AN ANALYSIS OF TIME-DEPENDENT TRANSCRIPTOMIC AND PHENOTYPIC CHANGES ASSOCIATED WITH REPAIR AND REGENERATION IN THE AIRWAY EPITHELIUM

BADRUL H. YAHAYA

This thesis is submitted as part of the course requirement for the degree of Doctor of Philosophy at the University of Edinburgh.

April 2010

DEDICATION

A special thank goes to my wife, Azila Abd Aziz for her endless love, encouragement and support for me to finish my PhD. To family in Malaysia especially my parents, Mr Yahaya Said and Mrs Som Lazim, for their continuous pray for my success and giving me a confident to finish my thesis. This PhD thesis I dedicated to my two little daughters, Balqis Batrisyia and Mya Safiya for their endless happiness which giving me strength and peace at all the times.

ACKNOWLEDGEMENT

My deepest appreciation is dedicated to my supervisors, Dr David Collie and Dr Gerry McLachlan for their great and untiring supervision, advices, support and trust during the course of this work. A special thanks also goes to Alison Baker and Peter Tenant (EBVC) for their assistant during a bronchial brushing, Dr Sionagh Smith (EBVC) for the advice and help in pathology, Dr Darren Shaw (EBVC) and Dr Caroline McCorquodale (RI) for their help for the statistical analysis, Dr Jeremy Brown (EBVC) and Dr Alan Pemberton (EBVC) for immunostaining, all staff in Cystic Fibrosis' Group for their endless support and help during my tough work in the lab and all individuals that contribute direct or indirectly during the course of my study. Last but not least my sponsors that support my study and the small grant for the research project, the Ministry of Higher Education of Malaysia and the Universiti Sains Malaysia, under the fellowship of Academic Staff Training Scheme (ASTS).

DECLARATION

This experiment and composition of this thesis are, unless otherwise stated, my own work. No part of this work has been, or will be, submitted for any other degree, diploma or qualification.

Badrul H. Yahaya (April 2010)

Contents

Abbreviations

AB-PAS	Alcian blue periodic acid staining
AEC	Alveolar epithelial cell
AJ	Adherens junctions
aRNA	Amplified ribonucleic acid
ATPase	adenosine triphosphatase
BASC	bronchioalveolar stem cell
BBr	Bronchial brushing
BLAST	Basic Local Alignment Search Tool
BP	Biological process
BrdU	5-bromo-2-deoxyuridine
BV	Blood vessel
CAM	Cell adhesion molecules
CC	Cellular component
CDC2	Cell division cycle 2
cDNA	Complementary deoxyribonucleic acid
CF	Cystic fibrosis
CFTR	Cystic fibrosis transmembrane conductance regulator
COPD	Chronic obstructive pulmonary disease
CTNNB	Catenin beta
CTNNBL1	catenin beta-like 1
DAB	Diaminobenzidine
DAVID	The Database for Annotation, Visualization and Integrated Discovery
DFCI	Dana-Farber Cancer Institute
DNA	Deoxyribonucleic acid
dNTPs	deoxyribonucleotide triphosphate
ECM	Extracellular matrix
EGFR	Epidermal growth factor receptor
EMBL-EBI	European Molecular Biology Laboratory-European Bioinformatics Institute
EMT	Epithelial-mesenchymal transition
Epi	Epithelium
EST	Expressed sequence tag
FDR	False discovery rate
GCs	Goblet cells
GO	Gene ontology
G-phase	Growth-phase
GSEA	Gene set enrichment analysis

HE	Haematoxylin & Eosin
IHC	Immunohistochemictry
IL	Interleukin
KEGG	Kyoto Encyclopedia of Genes and Genomes
KRT14	Keratin 14
KRT5	Keratin 5
LC	Left cranial
LCD	Left cranial diaphragmatic
LRC	label-retaining cells
LVD1	Left ventral diaphragmatic lobe 1
LVD2	Left ventral diaphragmatic lobe 2
MC	Mucosa
MEA	Modular enrichment analysis
MF	Molecular function
M-phase	Mitosis-phase
mRNA	Messenger ribonucleic acid
MYC	Myelocytomatosis
NCBI	National Center for Biotechnology Information
PCNA	Proliferating Cell Nuclear Antigen
QC	Quality control
RA	Right apical
RC	Right cranial
RCD	Right cranial diaphragmatic
REML	residual maximum likelihood
RIN	RNA Integrity Number
RNA	Ribonucleic acid
RT-PCR	Reverse transcriptase polymerase chain reaction
RVD1	Right ventral diaphragmatic lobe 1
RVD2	Right ventral diaphragmatic lobe 2
SCLC	Small cell lung carcinoma
SD	Standard deviation
SEA	Singular enrichment analysis
SM	Smooth muscle
SMC	Submucosa
SMG	Submucosal gland
SMOC2	Secreted modular calcium-binding protein 2
SP	Surfactant protein
S-phase	Synthesis-phase

TA	transient-amplifying
TAC	Transit-amplifying cell
TD	Terminal differentiated
TGF	Tumor growth factor
TLR	Toll-like receptor
TNF	Tumor necrosis factor
TUBB	Tubulin beta
UV	Ultra violet
WHO	World Health Organisation
BTS	British Thoracic Society
PBS	Phosphate buffered saline

ABSTRACT

CHAPTER 1

GENERAL INTRODUCTION

1.1 Mortality, morbidity and costs associated with respiratory disease

Whilst cardiovascular diseases are the leading cause of death in the world (causing almost 32% of all deaths in women and 27% in men in 2004), respiratory disease is also a major cause of mortality and morbidity (WHO, 2008). Indeed, as stated in the World Health Organisation (WHO) consultation on the development of a comprehensive approach for the prevention and control of chronic respiratory diseases (January 2001, Geneva), almost one-fifth of all deaths are attributable to respiratory illness, and both communicable respiratory diseases (tuberculosis, acute respiratory infection, pertussis, diphtheria, measles) and non communicable diseases (asthma, chronic obstructive pulmonary disease (COPD), bronchiectasis, cystic fibrosis, obliterative bronchiolitis, sarcoidosis) represent about 20% of the global burden of disease measured in disability-adjusted life years (DALYS) (WHO, 2001).

1.2 Global Health Challenges

Projections indicate that the four leading causes of death globally in 2030 will be ischaemic heart disease, cerebrovascular disease (stroke), chronic obstructive pulmonary disease and lower respiratory infections (mainly pneumonia) (WHO, 2008). Amongst all diseases, those of chronic process are of particular concern for future generations as a chronic disease epidemic is anticipated over the next two decades (Wang L et al., 2005, Srinath Reddy K et al., 2005).

In relation to current perspectives on chronic respiratory disease it is estimated that in the world today over 300 million people suffer from asthma, over 210 million from COPD and over 450 million from allergic rhinitis and other chronic respiratory diseases (Bousquet J and Khaltaev N, 2007).

These global observations and predictions are mirrored to a large extent in statistics pertaining to the United Kingdom. According to the report "The Burden of Lung Disease" produced by the British Thoracic Society in 2006 (BTS, 2006), 117,456 people died of respiratory disease in 2004 (compared to 106,081 dying of ischaemic heart disease). Respiratory disease kills one in five people in the UK. The financial burden associated with respiratory diseases is significant and increasing year on year. It cost the NHS £6.6 billion in 2004 (£3.0 billion in NHS primary care, secondary care and drugs), £1.9 billion in mortality costs and £1.7 billion in morbidity costs, as compared to £2,576 million in 2000 (BTS, 2001 , BTS, 2006).

Against this backdrop of impending epidemics and crippling cost forecasts, the WHO have launched, the Global Alliance against Chronic Respiratory Diseases. (GARD) - an organization dedicated to bringing together the combined knowledge of national and international organizations, institutions and agencies to improve the lives of millions of people affected by chronic respiratory diseases. Amongst the key objectives of the WHO strategy for prevention and control of chronic respiratory diseases are the need to improve secondary and tertiary prevention by identifying cost-effective interventions and optimised pharmacologic treatments.

1.3 Asthma and COPD: waves of inflammation, and a failure to repair

A common feature of both chronic asthma and COPD, the commonest chronic airway diseases, is the airway wall remodelling that occurs in association with cellular and structural changes and which, in both cases, leads to thickening, airway narrowing and airflow limitation. Whilst different cell types contribute to the pathology (asthma is characterized by an increase in CD4+ T cells and eosinophils whilst COPD is characterised by increased numbers of neutrophils and CD8+ T cells), it is recognised that repetitive bouts of inflammation and airway epithelial injury is a feature of both conditions.

In COPD, the repeated bouts of inflammation develop as a consequence of chronic inhalation of irritants. The inhaled irritant is usually tobacco smoke, but occupational

dust and environmental pollution are also invariably implicated. In asthma, inflammatory exacerbations are predominantly linked to viral infections, particularly rhinoviruses, whilst other causes including bacterial infection and exposure to allergens are also implicated. The structural and inflammatory features of COPD and asthma are described by (Jeffery PK, 2000, Jeffery PK, 2001). In asthma the epithelium appears to be more fragile than that of COPD, the epithelial reticular basement membrane (RBM) is significantly thicker and there is marked enlargement of the mass of bronchial smooth muscle. In COPD, there is epithelial mucous metaplasia, airway wall fibrosis, and inflammation associated with loss of surrounding alveolar attachments to the outer wall of small airways as well as an increase in bronchiolar smooth muscle (Jeffery PK, 2001, Szilasi M et al., 2006).

Whilst it is well accepted that a link exists between chronic repetitive bouts of inflammation and the development of airway wall remodelling in both conditions, questions do arise as to why there should be an apparent failure in airway epithelial repair in these settings? Answering these and other questions relating to the pathophysiology of chronic respiratory disease, demands an appreciation of the dynamic cellular and molecular events that normally operate to repair the airway epithelium following injury, and indeed, it is only through such appreciation that the abnormal process will be understood and candidate therapeutics potentially identified. It is likely that such approaches will increasingly revolve around 'omic' pathway and network analyses such that identification of candidate hub points as putative drug targets likely to influence the course of disease can be identified.

1.4 Developmental ontogeny and tissue repair

It is apparent that many of the pathways and molecules invoked as a consequence of tissue injury have their origins in developmental biology. Indeed, in bone and kidney, tissues with only limited repair potential, developmental pathways are co-opted during the healing process subsequent to injury (Imgrund M et al., 1999, Vortkamp A et al., 1998). Similarly in the skin, liver, pancreas, muscle, and vasculature, developmentally important genes have been linked to cellular repair (Beer HD et al., 2005, Steiling H and Werner S, 2003, Zhao P, 2004).

For the lung, although less information is available concerning the molecular regulation of cellular responses to injury and mechanisms of tissue repair, there is substantial literature available with respect to the developmental ontogeny of this organ (Cardoso WV and Lü J, 2006, Kimura J and Deutsch GH, 2007, Shannon MJ and Hyatt BA, 2004, Warburton D et al., 2005, Copland I and Post M, 2004). The development of the human lung is a complex process involving different stages and steps from the embryonic to post-natal periods. Shi et al., (2009) descibes the stages of respiratory system development from the formation of the lung and tracheal primodia from the primitive foregut, a formation of bronchial tree from lung epithelial tubules which then accompanied by development of the vascular structures (Shi W et al., 2009). Activation of specific developmental programs leads epithelial progenitor cells to differentiate into a variety of cell types such as secretory, ciliated, neuroendocrine, and basal cells (Shi W et al., 2009). Later, distal epithelial tubules undergo sacculation to form the primitive gas exchange region of the lung as well as differentiate to epithelial cell types I and II. Alveolar developed by the sub-division of primitive saccules by a process of septation called 'alveolization' (Shi W et al., 2009). This complex morphogenesis is controlled by the coordinated expression of transcription factors, growth factors, extracellular matrix molecules, integrins and intracellular adhesion molecules in local regulatory networks extending along the proximal to distal axis (Copland I and Post M, 2004, Demayo F et al., 2002, Warburton D et al., 2005). The formation of each component of the lung is regulated by the complex mixture of transcription factors. Transcription factors (such as GATA-binding protein 6 (GATA-6)) have the ability to influence lung epithelial specification whilst mesenchyme transcription factors (such as Gli2 and Gli3) exert influence over direct epithelial patterning (Copland I and Post M, 2004). The epithelial-mesenchymal interactions control the expression of signalling molecules during lung development and repair, and this control occurs in a temporal, spatial and cell-specific manner (Demayo F et al., 2002).

The mechanisms of cellular response and the control of transcription factors during early lung development and growth are gradually being revealed, as advantage is taken of mice with specific gene mutations, in order to question the specific role of

the respective gene products. Overall, the programmed regulation of the developmental process is undeniably complex in an organ that contains cells of at least 40 distinct lineages (Warburton D et al., 1998).

There is increasing interest, fuelled by genome-wide association studies, microarray profiling and animal disease models, in the role of developmentally regulated genes in complex chronic disease pathogenesis (Hedgehog, Secreted modular calcium-binding protein 2 (SMOC2), Notch, retinoic acid, transforming growth factor beta (TGF-b), SMAD) (Shi W et al., 2009). Clearly there is some merit to the concept that mechanisms of repair recapitulate elements of the developmental process (Demayo F et al., 2002, Warburton D et al., 2005). In relation to this concept, the disruption of any of these developmental pathways (Sonic Hedgehog, Notch, Retinoid, and TGF-beta) has also been found to influence the development of the post-natal lung, disease (Stick S, 2000) and COPD (Shi W et al., 2009, Shi W et al., 2007).

1.5 Models of airway injury and repair in the lung

In the context of characterising the processes occurring during healing subsequent to injury of the airway epithelium, advantage has been taken of a number of animal model systems which are broadly divided into those systems reliant on physical injury (Erjefält JS et al., 1995, Gordon RE and Lane BP, 1976, Hilding AC, 1965, Keenan KP et al., 1982a, Keenan KP et al., 1982b, Keenan KP et al., 1982c, Keenan KP et al., 1983, Lane BP and Gordon R, 1974, McDowell EM et al., 1979, Shimizu T et al., 1994, Wilhelm DL, 1953, Kim JS et al., 1997a) and those reliant on chemically induced injury – whether administered by inhalation or systemic means (Barrow RE et al., 1992, Barrow RE et al., 1993, Cox RA et al., 2005, Cox RA et al., 2009, Hoffman A et al., 2003, Hong KU et al., 2001, Hong KU et al., 2004b, Kim JS et al., 1997a, Kim JS et al., 1997b, Tuck SA et al., 2008). The nature and extent of injury varies widely between such model systems.

In this regard, Barrow RE et al., (1992) described four general categories of injury to the airway epithelium. The first is described as being reversible – healing occurs after the irritant is removed and the cells return to normal (Barrow RE et al., 1992).

The second involves exfoliation of individual cells but leaves the majority of the nonciliated columnar and basal cells intact (Barrow RE et al., 1992). The third involves desquamation of the group of cells, but leaves the basal cell layer intact (Barrow RE et al., 1992). The fourth category involves desquamation of cells, including the basal cells and after injury the repair process involves proliferation and migration of surviving nonciliated cells and differentiation to the normal phenotype, with the degree of proliferation reflecting the amount of injury that occurred (Barrow RE et al., 1992).

1.5.1 Chemical induction

Chemically-induced lung injury is well established as a means of achieving airway epithelial perturbation. Such techniques include exposing the lung to high concentrations of oxygen (O_2) or oxidant gases such as nitrogen dioxide (NO_2) or ozone (O_3). Cell death from hyperoxic injury may occur through either apoptotic or non-apoptotic pathways, possibly via free oxygen radicals (Liu X et al., 2006). It is appreciated that the different types of cell present at the air interface in the lung exhibit a range of sensitivity to the effects of chemical exposure. For instance, the alveolar epithelial cell type I (AEC I) is particularly sensitive to bleomycin and the first to be injured in the context of exposure to this chemical whereas alveolar epithelial cell type II (AEC II) have a more variable sensitivity (Izbicki G et al., 2002). Subsets of Clara cells in the conducting airway epithelium are recognised to be differentially sensitive to naphthalene (Giangreco A et al., 2002, Hong KU et al., 2001). These susceptibilities should be borne in mind when interpreting the response to injury mediated by these chemicals.

Whilst the general utility of chemically-induced lung injury in small animal models is reflected in their popularity it is important to reflect upon their comparative relevance. Indeed, naphthalene is a cytochrome P-450-activated Clara cell cytotoxicant that, in mice, causes swelling and vacuolation of Clara cells in terminal bronchioles leading to exfoliation and necrosis of the majority of cells in the airways of the deep lung (Plopper CG et al., 1992). Mice appear to be much more susceptible to this toxicant than other species, including nonhuman primates, rats and hamsters

(Plopper CG and Hyde DM, 2008) and this fact, coupled with the predominance of this cell type in the proximal conducting airway in mice relative to humans suggests that potential exists for mechanisms of repair evoked as a consequence of such injury to vary between species. Indeed, the cellular composition of the tracheal and tracheobrochial tree of the airway does differ between species – with Clara cells in mice representing 49% of the total epithelial cells in tracheal epithelium whilst at the same site in monkeys or humans less than 1% of the cells are Clara cells (Plopper CG and Hyde DM, 2008). Further, sub mucosal glands are important secretory glands that are present in the major airways and bronchioles of humans and whilst in mice the structure, cellular composition, and density of tracheal submucosal gland (SMG) are similar to those seen in humans, this is in fact the only site where they are seen in this species (Innes BA and Dorin JR, 2001). Lastly, there are obvious differences between large and small animals in the thickness of the epithelium with the former being characterised by a thicker epithelium. These differences raise the possibility that mechanisms of repair evoked as a consequence of injury may also vary between species (Plopper CG and Hyde DM, 2008).

Chemically-induced lung injury normally causes relatively mild injury involving shedding of the epithelium whilst leaving attenuated squamous basal cell layer and underlying tissue architecture (Park KS et al., 2006). Any remaining or lesion-bordering epithelial cells have the capacity to quickly regenerate a pseudostratified epithelium and both basal and/or nonciliated columnar epithelial cells have the potential to serve as primary progenitor cells in this regard (Rock JR et al., 2009, Rawlins EL et al., 2008). Although this technique allows for perturbation of the airway epithelial morphology, the extent and location of the epithelial damage is non-uniform throughout the exposed areas.

Such animal models coupled with transgenic and cell type-specific knockout strategies have been used to probe the importance of beta-catenin regulated gene expression in bronchiolar regeneration and repair on this species (Zemke AC et al., 2009). Manipulation of gene-specific cell types also promises the ability to trace the roles of each epithelial cell (i.e Clara and alveolar type I and II) in airway epithelial

proliferation and differentiation during lung development. This includes studies directed towards determining the role(s) of GATA-6 in regulating bronchioalveolar stem cells (BASCs) during development and airway regeneration (Zhang Y et al., 2008).

Such chemically-induced models of lung injury are not confined to the domain of small animal models. For instance a model of smoke-induced lung injury in sheep is widely considered of appropriate comparative relevance to the human (Alpard SK et al., 2000, Wang CZ et al., 1990). With similarities in terms of the immediate response to direct lung injury, the subsequent inflammatory response syndrome and similar therapeutic opportunities and application including mechanical ventilation (with associated risk of baro/volutrauma) the model continues to inform. In addition, the use of the goat as a model for smoke induced lung injury was reported as reliable and of comparative relevance particularly in the context of histopathological changes in the tracheobronchial tree (Walker HL et al., 1981). One such study demonstrated that within 1 day of exposure to toxic smoke the columnar epithelium sloughed intact from the trachea with a concomitant reduction of nearly 35% in the basal cell population. Complete repair occurred within 18-22 days following injury (Barrow RE et al., 1992)

It is however difficult to align such studies across species wherein the extent of the injury itself, and presumably the evoked mechanisms of repair are a prima facie reflection of fundamental differences in anatomy, dosimetry and xenobiotic metabolism.

1.5.2 Physical methods: mechanical trauma

Whilst not immune from the effects of species differences in the anatomy of the tracheobronchial tree, methods that use physical means to perturb the airway epithelium have provided an alternative strategy to study the mechanisms involved in successful repair and regeneration following injury. Table 1.1 shows a summary of the mechanical injury techniques used in small animal models systems. The extent of

the injury varied - in some models no epithelial cells remained on the denuded epithelium, whilst in others the basal cell layer was left intact.

Wilhelm DL (1953) was the first to apply a mechanical injury, induced by curettage, to the tracheal airway mucosa in a small animal (rat) model system. The induced injury was relatively severe – involving the complete removal of the intact mucosa (Wilhelm DL, 1953). A subsequent study carried out by Hilding AC (1965) demonstrated the use of cotton swabs to mechanically injure the trachea of calves. In this study, the lesion was produced by passing a weighted cotton swab over the tracheal epithelium and caused only the exfoliation and loss of mature columnar cells, leaving the basal layer intact (Hilding AC, 1965).

A series of studies carried out by Keenan KP and colleagues studied the process of repair and regeneration following mechanical injury applied to the tracheae of hamsters. In these experiments, the ventral surface of the trachea was exposed surgically and incised at the intercartilaginous portion immediately posterior to the larynx. A mechanical injury was induced by a single stroke of a small stainless steel probe along the ventral quadrant of the tracheal mucosa from the carina to the larynx. The basal layer was left intact (Keenan KP et al., 1982a, Keenan KP et al., 1982b, Keenan KP et al., 1982c). These studies were similar in principle to a study carried out by Lane BP and Gordon R (1974), in which the dorsal surface of the trachea was lightly stroked by a blunt probe. In this latter study, there was loss of most of the damaged cells in the area of injury leaving a single discontinuous layer of basal cells with occasional residual ciliated or goblet cells (Lane BP and Gordon R, 1974).

In a study of tracheal epithelial regeneration following a mechanical injury in hamsters (McDowell EM et al., 1979), a severe injury was induced by scraping with a blunt probe. This method lethally damaged all the epithelial cells in focal areas in order to maximise the extent of the regenerative response and to induce the involvement of all cell types during airway epithelial regeneration and repair (McDowell EM et al., 1979). De-differentiation and migration of the epithelial cells bordering the lesion were clearly described in this study.

Similar processes were studied in rats following a mechanical injury (Shimizu T et al., 1994). The epithelium was lethally injured by exposing the ventral surface of the tracheal by surgical incision at the intracartilaginous portion below the larynx. A stainless steel probe was inserted through the surgical incision and stroked the ventral surface of the trachea from carina to larynx (Shimizu T et al., 1994). No viable cells were present at the area of damage. The transition area between damaged and undamaged epithelium was abrupt with normal and viable epithelial cells seen at the margins of the wounded area. Those cells from the wound margins appeared to flatten and migrate across the wound site to cover the damaged epithelium. Thereafter the migrated cells contributed towards the formation of an initially poorly differentiated multilayered epithelium (Shimizu T et al., 1994).

In a study carried out by Erjefalt JS et al., (1995) involving *in vivo* restitution of guinea-pig airway epithelium, the epithelium was removed along the dorsal aspect of the trachea by a steel-probe. This technique did not cause damage to the underlying, denuded basement membrane. Basal cells remained attached to the basement membrane. It was reported in this study that the epithelial cells at the wound edge flattened and migrated over the denuded basement membrane within 15 minutes. Cells underwent dedifferentiation with ciliated cells losing their cilia and secretory cells displaying an intense granule discharge at the wound edge (Erjefält JS et al., 1995).

Models	Organ	Tools	Effect on epithelium	Severity	References
Rat	Tracheae	Curettage	Complete removal of intact mucosa	Severe	(Wilhelm DL, 1953)
Calve	Tracheae	Cotton swab	Basal layer retained	Mild	(Hilding AC, 1965)
Hamster	Tracheae	Blunt probe	Basal layer retained	Mild	(Keenan KP et al., 1982a, Keenan KP et al., 1982b, Keenan KP et al., 1982c, Keenan KP et al., 1983)
Rat	Tracheae	Blunt probe	Basal layer retained	Mild	(Lane BP and Gordon R, 1974)
Hamster	Tracheae	Blunt probe	complete removal of intact mucosa	Severe	(McDowell EM et al., 1979)
Rat	Tracheae	Blunt probe	complete removal of intact mucosa	Severe	(Shimizu T et al., 1994)
Guinea-pig	Tracheae	Steel-probe	Scattered basal cells remained attached to the BM	Mild	(Erjefält JS et al., 1995)
Guinea-pig	Tracheae	Flattened scrapper	Complete removal of intact mucosa	Severe	(Kim JS et al., 1997a)

Table 1.1: Small animal models involving mechanical injury to the trachea
Variation in the severity of injury can be expected to influence the timing of the repair process

1.6 Patterns of repair in the airways

1.6.1 Dedifferentiation, migration, proliferation and re-differentiation

The process of airway epithelial repair subsequent to severe injury appears to follow well-defined stages. At an early juncture, unaffected epithelial cells bordering the lesion migrate to cover the damaged area. These migrated cells appear to form a multilayered epithelium, a hyperplastic epidermoid metaplasia. Lastly, these cells re-differentiate to re-establish a normal pseudostratified epithelium (McDowell EM et al., 1979). Subsequent to mild injury in which viable basal cells are left at the surface of the damaged area, regeneration is completed by epithelial hyperplasia and epidermoid metaplasia, followed by differentiation (Lane BP and Gordon R, 1974, Hilding AC, 1965). In this latter instance, basal cells were reported to be the primary progenitor cells contributing to the cellular migration and proliferation in response to injury, whereas in the case of more severe injury (with no basal cells left intact), the secretory and nonciliated cells (goblet cell in particular) were thought to be the primary progenitor cells involved in the cellular de-differentiation, migration and re-differentiation.

It is therefore concluded that the common features that seem to characterize airway epithelial repair following physical injury induced by mechanical means involve the dedifferentiation of cells bordering the lesion, migration of the flattened cells over the wound area, proliferation and re-differentiation. In particular, various cell types, including basal, ciliated and secretory cells, appear to possess the capacity to dedifferentiate, to migrate, and to either re-differentiate, or trans-differentiate to give rise to other types of cells. Indeed secretory cells have been shown to dedifferentiate and become flattened epithelial cells when seeded in normal tracheal epithelial cell culture and thereafter re-differentiate to normal morphology whilst basal cells were less frequently observed in a similar process (McDowell EM et al., 1987). These observations further suggest that potential may exist to manipulate conditions following injury or even in the context of airway disease such that the airway epithelium can be restored to its optimum.

1.6.2 The role of stem/progenitor cells

As previously expressed, many of the mechanisms that normally operate to engineer tissue repair and which are overwhelmed in the prelude to lung pathology are believed to have their foundation in lung development. A further aspect of this corollary is that during development, self-renewing tissues are imbued with resident, tissue-specific stem cells, so-called adult somatic stem cells (Neuringer IP and Randell SH, 2004). These stem cells are a subset of undifferentiated cells with the capacity to maintain multipotency in the context of the physiologic domain in which they reside. In this context, adult lung stem cells are capable of abundant self-renewal and regeneration, and should act as stem/progenitor cells in response to injury and effect local repair (Emura, 2002, Engelhardt JF, 2001, Liu X et al., 2006, Otto, 2002, Stripp and Shapiro, 2006). As such, these cells may serve as a viable target to manipulate this process in the context of the abnormal repair patterns that threaten normal lung physiology. Several cell types of the lung capable of functioning as stem/progenitor cells in response to injury have been identified; these cells are thought to localize to proximal airway sub mucosal gland ducts, intercartilagenous ring regions, neuroepithelial bodies, and terminal bronchioles/bronchoalveolar duct junctions (Adam Giangreco et al., 2002, Carla F. Bender Kim et al., 2005, Duncan W. Borthwick et al., 2001, Hong KU et al., 2001). The cells identified as progenitor or stem cells in the lung appear to vary according to the lung compartment (Berns A, 2005, Bishop AE, 2004, Emura M, 2002, Engelhardt JF, 2001, Neuringer IP and Randell SH, 2004, Otto WR, 2002). Equally these observations potentially imply that the specific repair mechanisms evoked in response to lung injury will draw upon several sources of stem/progenitor cells according to the nature and extent of the damage. In this regard it has been suggested that slight or moderate injury will result in resident progenitor cell activation to restore tissue homeostasis whereas severe injury and extensive epithelial cell loss will promote stem-mediated repair (Giangreco A et al., 2009).

Proliferative potential is often taken as a defining characteristic for stem cells (Stewart, 2006). The normally slow rate of lung epithelial turnover is easily accommodated by division, migration and differentiation of resident and/or mobile

stem/progenitor cells (Raiser DM et al., 2008). Stripp BR and Reynolds SD (2008) have described the terminology for the progenitor cells as any cell that has the capacity to proliferate and this terminology is commonly used to indicate a cell that is in the process of cell division or has the potential to enter the cell cycle (Stripp BR and Reynolds SD, 2008). In contrast to other tissues such as the gut and skin where the epithelium renews on a frequent basis, murine lung epithelium has little or no turnover in the absence of exogenous stimuli (Raiser DM et al., 2008). Therefore, under such conditions, it is difficult to capture the nature of the specific events involved in renewal, and it is only when normal tissue homeostasis is stressed by the need to engineer widespread tissue repair that the relative importance of different cell types within the tissue can be ascertained. It is possible that lung stem cell populations may be a subset of the same cells that function daily as contributors to the gas-exchange machinery, and their stem cell characteristics may only be realised in times when injury of more specialized, or more differentiated, cells occurs (Raiser DM et al., 2008). For this reason, several lung injury models have been developed to probe the mechanisms involved in tissue inflammation and repair, and will likely continue as the most useful approaches to identifying stem cell populations and understanding the regulation of stem cell fates in the lung (Liu X et al., 2006). These include radiation (Theise ND et al., 2002), sulphur dioxide (SO_2), polidocanol (Duncan W. Borthwick et al., 2001), bleomycin (Izbicki G et al., 2002) and naphthalene (Carla F. Bender Kim et al., 2005) as well as the previously discussed mechanically-induced lung injury (Keenan KP et al., 1982a, Keenan KP et al., 1982b, Keenan KP et al., 1982c, Keenan KP et al., 1983, Kennedy AR et al., 1978, Lane BP and Gordon R, 1974, Lane BP and Gordon RE, 1979, McDowell EM et al., 1979, Shimizu T et al., 1994, Wilhelm DL, 1953).

The maintenance and fate of the stem cells populations are regulated by their local anatomical and chemical microenvironment, or niche. In order to maintain stem cell phenotypes and thus promoting appropriate cell fate and migration activities, niches discrete microenvironments of specialized cell types, matrix, and diffusible factors such as cytokines and growth factors (Liu X et al., 2006, Watt FM and Hogan BL, 2000). Because stem cells divide infrequently under normal conditions DNA-

labelling techniques with detectable nucleotide analogues have been employed to track stem cells *in situ*. BrdU (5-bromo-2-deoxyuridine) is one such common nucleotide analogue that is classically used to track label-retaining cells (LRCs) after a prolonged "washout" period that dilutes the label within the more rapidly cycling transient-amplifying (TA) cells – committed progenitors derived from stem cells whose propensity for rapid cycling has a defined capacity.

1.7 Chronic airway disease – cellular mechanisms

One frequently observed pulmonary response to injury (such as acute lung injury (ALI) and acute respiratory distress syndrome (ARDS)) is the activation of endothelial cells and macrophages (alveolar and interstitial), up regulation of adhesion molecules, and the production of cytokines and chemokines. The activation of these activities induces a massive sequestration of neutrophils within the pulmonary microvasculature. Transmigration of these cells into the alveolar space which then release a variety of cytotoxic and proinflammatory compounds including proteolytic enzymes, reactive oxygen species (ROS) and nitrogen species, cationic proteins, lipid mediators, and additional inflammatory cytokines (Chow CW et al., 2003).

Persistent activation and chronic secretion of a variety of proinflammatory epithelial factors resulted from a chronic injury or defective repair of the surface epithelium which include epithelial-derived growth factor and granulocyte-macrophage colony stimulating factor which act to induce alterations to the epithelial reticular basement membrane, via activation of adjacent fibroblasts and myofibroblasts (Jeffery PK, 2004). Injury to the superficial epithelium is normally accompanied by mitotic activity in the remaining cells and rapid restoration of the denuded surface (Jeffery PK, 2001). However, whilst it is still unclear as to exactly how these mechanisms and factors fail to ensure continuous repair in the context of the disease state the assumption does arise that the repeated injury scenario eventually leads to defects in the normal repair pathways.

As previously stated, the normal response of the airway surface epithelium to injury includes a succession of cellular events, varying from the loss of the surface epithelium integrity to partial shedding of the epithelium, or even complete denudation of the basement membrane (Puchelle E et al., 2006). The repair of the epithelium includes spreading and migration of basal and secretory cells, followed by proliferation and differentiation of epithelial cells (Puchelle E et al., 2006). In the event that these processes fail to adequately repair the airway then pathology and remodelling may ensue with subsequent pathophysiological impact. Indeed, the development of squamous metaplasia and mucous hyperplasia in COPD as part of remodelling process are considered to disturb the innate immune functions of the airway epithelium (Puchelle E et al., 2006). The airway epithelial remodelling in COPD may induce an abnormal localization and expression of ion channel proteins such as cystic fibrosis transmembrane conductance regulator (CFTR), which may cause abnormal epithelial functions, such as periciliary epithelial dehydration, as observed in cystic fibrosis (CF) due to the defective CFTR protein (Puchelle E et al., 2006). A hyposecretion of mucus in the submucosal gland could also be responsible for a decreased release in the airway lumen of glycoproteins, proteins, and peptides which normally represent the biochemical barrier against microorganisms (Puchelle E et al., 2006).

1.8 Tools and resources to study airway epithelial repair

The complexity of the airway wall tissue architecture, the multiple cell lineages involved and their physical, autocrine and paracrine interactions has contributed to difficulties and delay in understanding the mechanisms involved during airway regeneration and repair. Accepting the likelihood that cells, genes and regulatory pathways involved during lung morphogenesis and development may also play a crucial role during lung injury and repair it is conceivable that key default pathways do indeed exist and that these pathways are evoked in response to acute injury. Further, it is only when such pathways are perturbed or overwhelmed that the sequences of events result in pathology. To get a wider perspective of the genes involved during airway repair, global gene expression profiling has emerged as the ultimate investigators' tool in that it has the capability to screen the expression level

of several thousands of genes in each individual sample. Such screening raises the potential that individual genes or gene regulatory pathways with key roles in the repair process can be identified and, in future, that such identification will lead to more informed therapeutic targeting in the context of complex chronic lung diseases. Whilst the past decade has seen the development of such high throughput methods to collect proteomic, genomic and transcriptomic data, and the increasing application of systems biology approaches to understand biological processes there has, to our knowledge, been limited attempts to marry these resources to the well characterized process of airway epithelial repair following physical injury.

The use of this high-throughput technology has been fundamental to identifying genes with key roles in the progression of diseases such as asthma (Aoki T et al., 2009, Tölgyesi G et al., 2009), lung cancer (Rohrbeck A et al., 2008) and COPD (Bhattacharya S and Mariani TJ, 2009) and the promise remains that new biomarkers will be detected and diseases better understood through sub-classification according to their underlying mechanisms (Bhattacharya S and Mariani TJ, 2009) – in turn adding leverage to genome-wide disease association studies.

Gene expression profiling continues to play a vital role in the context of understanding the response of animal model systems to experimental manipulation. Such studies have encompassed the response of the lung to injury. As a recent example, in a study of gene expression profiling in the lungs of wild-type mice exposed to naphthalene at various time points, 3,552 genes were significantly up-regulated and 1,199 genes significantly down-regulated following injury (Snyder JC et al., 2009). Analysis of gene expression in the repairing lung revealed prominent clusters of up-regulated genes with putative roles in regulation of the extracellular matrix and cellular proliferation (Snyder JC et al., 2009). It was also suggested that the extracellular matrix (ECM) is dynamically remodelled in response to selective epithelial cell injury and that this process is activated without resolution in the setting of defective airway epithelial repair (Snyder JC et al., 2009). Other examples of such expression profiling in the context of airway epithelial injury in small animals relate to assessment of the alterations in the transcriptional program of the lung in a mouse

model of acute lung injury (ALI) induced by *E. coli* lipopolysaccharide (LPS) (Jeyaseelan S et al., 2004), acute respiratory allergen exposure of transgenic BALB/c mice by a single intratracheal aspiration exposure to Metarhizium anisopliae crude antigen (MACA) (Pucheu-Haston CM et al., 2010), the gene expression profile of mice exposed to cigarette smoke (Cavarra E et al., 2009) and co-exposure of mice to cigarette smoke, lipopolysaccharide, or smoke plus lipopolysaccharide by inhalation (Meng QR et al., 2006).

In a clinical study, using fiberoptic bronchoscopy and brushing to sample the human bronchial epithelium, Harvey et al., (2007) investigated the modification of gene expression in response to cigarette smoking. Such transcriptomic screening revealed that in comparing between phenotypically-normal smokers (no symptoms, normal lung function, normal imaging) and non-smokers, microarray analysis of gene expression of the small airway epithelium demonstrated up- and down-regulation of genes in multiple categories relevant to the pathogenesis of COPD, including genes coding for cytokines/innate immunity, apoptosis, mucin, response to oxidants and xenobiotics, and general cellular processes (Harvey BG et al., 2007). Such studies highlight the depth of information that can exist behind the screen of apparently similar phenotypes and augers well for the future application of such technology in predicting disease onset and hence directing appropriate prophylactic and/or therapeutic strategies.

Another clinical study that relates to our interest in understanding the molecular regulation of gene expression during airway regeneration and repair is a study carried out by Heguy A et al., (2007) to identify genes participating in the airway epithelial repair of healthy volunteers after exposure to mechanical injury induced by bronchial brushing via fiberoptic bronchoscopy. This study revealed that 1,196 genes were significantly differentially expressed at day 7 post injury when compared with day 0 (un-injured) (Heguy A et al., 2007). Among these genes, 60% were associated with the G_2 and M phases of the cell cycle, and it was suggested that this reflected the fact that most of the cells at this stage of repair had entered the late stage of cell division.

1.9 Large animal models

The aforementioned clinical studies in particular emphasise the value of being able to accord gene expression profiles with some experimental state – whether by virtue of exposure to cigarette smoke, or bronchial brushing injury – and equally suggest the value of being able to measure and adjust for the influence of between- and within-subject variability in gene expression in analysing experimental effects. Naturally there are limits to the nature and extent of intervention possible in a clinical research setting and to yield information pertinent to more severe pathology it is necessary to either collect data from relatively poorly controlled clinical case material and/or from relevant animal model systems. In the latter regard it is important that the system selected is indeed of comparative relevance, offers the same facility to study expression profiles in selected cells over time in the same subject and is backed up by the necessary reagents and tools to allow a complete characterisation of the process in question.

The issue of comparative relevance is hotly debated by every proponent of every animal model system (Abraham WM, 2008, Bates JH et al., 2009 , Canning BJ and Wright JL, 2008, Chapman RW, 2008, Keir S and Page C, 2008, Kumar RK and Foster PS, 2001) and suffice is to say that no particular system is ideal for all situations and that each system generally has meritable comparative aspects, whether in the context of normal lung structure or function, or in relation to response to injury. In relation to the second requirement, that of following expression profiles in selected cell populations (available for harvesting during routine bronchoscopic procedures) in the same individual animals over time it is clear that such facility is only really provided by larger animals. Whilst it is the last requirement that has largely fuelled the popularity of the mouse as a model system for investigating the pathogenesis of airway disease – it is inevitable that, with the inexorable progress underway in relation to genome characterisation across species, the nature and extent of tools and technologies available for larger animals will similarly increase.
Whilst some large animals have been employed to study the dynamics of repair in the tracheal epithelium, notably in the context of either very mild trauma (in calves (Hilding AC, 1965)) or induced full thickness mucosal defects (in dogs (Koempel JA

et al., 1998) and rabbits (Hardillo J et al., 2001)), no studies have, to our knowledge, sought to define the patterns of repair that follow injury to the bronchial epithelium in large animals. Previous studies within our group have indicated the utility and comparative relevance of the sheep as an appropriate large animal model system to study lung disease pathogenesis and follow the response to lung-directed gene delivery (McLachlan G et al., 2007). However, to the best of our knowledge, there have been no previous studies directed towards the transcriptomic profiling of sheep airway epithelial gene expression following lung injury. Such studies would supplement the recently published transcriptional data in relation to repair following bronchial brush biopsy in human subjects (Heguy A et al., 2007) and serve as a useful comparison of the respective processes underlying repair in a relevant model system.

1.10 The case for the sheep as a model of comparative relevance

In considering the most appropriate model system to bring together the time-dependency, the injury itself, the requirement for directly sampling the injured sites themselves and the potential wider relevance to human disease we elected to use the sheep as our system of choice. Sheep are docile, easily handled animals that are readily purchased from commercial stock markets. Housing and maintenance costs obviously depend on the specific research requirements, however these certainly compare favourably with other species frequently used as respiratory models, such as dogs or primates. Sheep tolerate routine procedures such as blood sampling very well and do not constitute a particular anaesthetic risk. For these and other reasons the sheep is favoured as an experimental animal in surgical science (Borrie J and Mitchell RM, 1960). With specific respect to the respiratory system, sheep lungs are anatomically closer to those of man than are dog, cat or monkey lungs (McLaughlin RF et al., 1994) and when compared to dogs, sheep appear to be better suited as models for human gas exchange (Werlen C et al., 1984). The size of sheep lungs facilitates the application of various techniques available for the analysis of human cardiorespiratory function. Further in sheep, time sequential analysis of the respiratory tract, including physiological, biopsy and lavage procedures, is well tolerated and can be repeated at fortnightly intervals with no procedure-related

effects (Bégin R et al., 1981). Thus, dynamic events following natural or experimental respiratory intervention can be followed in this species.

Prior to the study of Halmagyi DF and Colebatch HJ, (1961b) negligible data existed in the literature regarding cardiorespiratory function in sheep. The data reported by these authors included tidal volume, gas exchange, pulmonary mechanics and blood gas and acid-base values for normal anaesthetized sheep (Halmagyi DF and Colebatch HJ, 1961b). These authors subsequently went on to utilize their experience of lung function assessment in sheep to examine cardiopulmonary dysfunction in sheep in response to various interventions of both physiological and applied interest (Colebatch HJ and Halmagyi DF, 1961, Colebatch HJ and Halmagyi DF, 1963, Halmagyi DF and Colebatch HJ, 1961a). This approach, whereby measurement protocols and techniques for lung function analysis are developed and refined prior to their use in the study of respiratory disease, has also been followed by Bégin and colleagues in an evaluation of asbestosis in sheep (Bégin R et al., 1983a, Bégin R et al., 1983b). Other groups have developed and used lung function analysis techniques in the sheep to examine the airway (Abraham WM et al., 1983, Abraham WM et al., 1981, Ahmed T et al., 1980, Hutchison AA et al., 1984, Hutchison AA et al., 1982, Wagner EM and Mitzner WA, 1990, Wanner A et al., 1979, Wanner A and Reinhart ME, 1978), lung inflammatory responses (Wheeler AP et al., 1990a, Wheeler AP et al., 1992, Wheeler AP et al., 1990b), lung function during anaesthesia (Coulson NM et al., 1989, Coulson NM et al., 1991), gas exchange of the lung (McNeil JS et al., 1991, Werlen C et al., 1984) and physiological responses to exercise (Mundie TG et al., 1991). Further, and beyond the experience of our own group, the sheep is favoured as a comparative model in respiratory research with examples of applications ranging from smoke inhalation injury (Barrow RE et al., 1992, Barrow RE et al., 1993, Cox RA et al., 2005, Cox RA et al., 2009), to the adult respiratory distress syndrome (Meyrick B, 1986), asthma (Kirschvink N and Reinhold P, 2008) and emphysema (Hoffman A et al., 2003). The sheep and hamster have a large number of airway epithelial cell types which secrete acid glycoprotein in the proximal airways and the sheep has glands in abundance (Mariassy AT and Plopper CG, 1983). With more basal and ciliated cells than small animals, the respiratory

epithelium of the sheep is 4-5 times the thickness of that reported for hamsters and rats (8.6 microns in the hamster to 56.8 microns in the sheep) (Evans MJ et al., 1989) and relatively similar to the height of the epithelium reported in human (40 - 50 microns) (Plopper CG and Hyde DM, 2008).

As our model of injury we elected to employ the clinical technique of bronchial brushing to selectively damage small areas of epithelium and thereby stimulate repair processes in those areas. This technique is routinely used within our group to selectively harvest small populations of airway epithelial cells from sheep and prior experience (D Collie, personal communication) indicates that the technique is both well tolerated and resolves rapidly within a period of 1-2 weeks, at least on a gross level.

1.11 AIMS OF THE THESIS

Airway epithelial damage is a prominent feature of chronic airway disease in the lung. Whilst a number of classical studies have elegantly characterised the response of the airway epithelium to mechanical trauma at the cellular and tissue structural levels a gap currently exists in relation to characterising and describing the transcriptional features that parallel the repair process. The recent description of transcriptional events associated with the repair process following bronchial brush biopsy in human subjects provided an opportunity to quantify the effects of similarly induced injury in a large animal model of proven comparative relevance for respiratory research.

The objectives that this thesis sets out to achieve are stated thus:

1. To develop a bronchial brushing technique as a model system of lung injury and repair in sheep.
2. To investigate the time course-effect of the histopathological changes of the airways in response to physical injury
3. To study the effect of the bronchial brushing at cellular and molecular levels using histopathological techniques and large scale transcriptional analysis using microarray technology

CHAPTER 2

A PRELIMINARY ANALYSIS OF TIME DEPENDENT CHANGES ASSOCIATED WITH AIRWAY EPITHELIAL REPAIR IN SHEEP

2.1 INTRODUCTION

The lung is a relatively stable organ with low rates of cell turnover, particularly in the airways. Indeed, with less than 1% of epithelial cells proliferating at any given point in time (Demoly P et al., 1994, Boers JE et al., 1996), there appears to be little need for local self-renewal under normal circumstances. However, following airway epithelial injury it is clear that the epithelium has great capacity to repair and that the mechanisms involved normally allow for the total restitution of the original complex structure (Erjefält JS et al., 1995, Erjefält JS and Persson CG, 1997). Understanding the mechanisms underlying the process of regeneration and repair of airway epithelial structures demands close characterisation of the associated cellular and molecular events. In the presence of acute epithelial injury, it is recognised that cell proliferation is a feature of repair responses that lead to resolution of the injury. Indeed, evidence suggests that there is rapid repair and resolution following physical injury, at least in small animal models (Keenan KP et al., 1982a, Keenan KP et al., 1982b, Keenan KP et al., 1982c, Shimizu T et al., 1994, Kim JS et al., 1997a, McDowell EM et al., 1987, Wilhelm DL, 1953). As detailed in the introductory chapter to this thesis (Chapter 1; Sections 1.5.1 and 1.5.2) such models indicate that an initial key event is the propensity of the cells bordering the lesion to dedifferentiate and flatten. These flattened cells then appear to migrate inwards and over the denuded area to restore the barrier function of the epithelium. Proliferation is also a key feature of complex and chronic lung disease. For example, in asthma, airway epithelial cell proliferation varies directly with disease severity (Cohen L et al., 2007) whereas the same relationship does not appear to hold for COPD (Miller YE et al., 2007).

In many disease conditions, such as asthma and COPD, where the pathogenesis is believed to involve chronic repetitive airway epithelial injury and inflammation, the damage fails to fully resolve and instead results in tissue remodelling and pathology which may interfere with function (Jeffery PK, 2001). Key to understanding the factors that divert repair away from resolution and towards remodelling is a close understanding of the dynamic events that operate to resolve epithelial injury.

Whilst virtually all *in vivo* physical injury models of airway epithelial repair are small animal based, it is at least debatable whether these systems offer valid comparison with similar injury experienced in the human lung as such models have not been widely studied in that context. Indeed, it would appear that only one such model has been examined in any detail (Heguy A et al., 2007) and, whilst apparent opportunity exists for comparison, it is notable that the time points examined in this study (days 7 and 14 after injury) extend beyond the timing of the majority of changes in architecture and cellular dynamics that are well characterised in small animal systems ie by day seven the epithelium has regained a normal appearance. In contrast, in the study of (Heguy A et al., 2007) in human subjects, the morphological abnormality of the injured epithelium 7 days after injury was associated with significant changes in gene expression profile – indicating that healing was still underway and normality, in both senses, was achieved by 14 days after injury.

This contrast may reflect underlying species differences in the temporal dynamics of the repair process or in the underlying structure and composition of the airway epithelium however inherent differences in the extent of injury may also contribute. This is unlikely to be resolved until some of these variables can be aligned between studies and questions probed at a deeper mechanistic level.

To this end one useful contribution would be to develop model systems with closer relevance to the human lung – both in the context of providing a closer analogue of human airway wall architecture, but also in the context of potentially aligning perturbation and assay techniques.

As far as structural considerations are concerned the human airway epithelium is characterised as pseudostratified columnar and includes ciliated, secretory and parabasal cells that are linked to a foundation of basal cells that anchors the epithelium to the lamina reticularis (Harkema JR et al., 1991). The proximal airways of small mammals, on the other hand, typically are only one or two cells thick and those cells rest on only a very sparse network of basal cells (Evans MJ et al., 1990, Evans MJ et al., 1989). Notably in the sheep, Mariassy AT and Plopper CG, (1983)

quantified eight major cell groups in sheep airway based on their differences in cell morphology, staining properties, and distribution. These cells were mucous cell categories (M1, M2, M3, and M4), ciliated, basal, Clara cells, and serous cells, with the latter being restricted to only the submucosal gland areas (Mariassy AT and Plopper CG, 1983). At the level of the distal airways Clara and ciliated cells predominated in the lining epithelium whereas more proximally mucous (M1, M2, M3, and M4), basal, and Clara cells shared the epithelial lining (Mariassy AT and Plopper CG, 1983). Whilst it is likely that the range of cell types present in the intact epithelium and their properties will have a role in determining the response to injury, there have been limited studies to date that have attempted to accurately define such responses in the context of a large mammalian lung *in vivo* (Barrow RE et al., 1992, Barrow RE et al., 1993, Hilding AC, 1965).

In terms of aligning protocols, since bronchoscopic procedures form the basis of clinical respiratory research, a significant advantage could be gained from animal model systems that can offer the same facility of access, together with the potential of repeated assessment within the same individual.

In identifying the potential of a novel large animal model system to study the dynamics of airway epithelial repair following physical injury induced by mechanical means we sought to develop the sheep as the animal in question. The advantages of using sheep as our animal model system for lung injury and repair have previously been described (Chapter 1; Section 1.10).

In deference to the human literature available at the time, induction of injury through equivalent employment of a bronchial brushing technique was selected – a routine and safe method for cell sampling in clinical and veterinary clinical respiratory medicine. In addition to standard means of morphological assessment and recognising the key role that proliferation plays during repair and resolution following injury we sought to characterise this aspect of the response by BrdU labelling. In addition, we also sought to determine the influence of injury on mucus-producing cells in the epithelium as an indicator of airway differentiation.

2.2 OBJECTIVES

The objectives of the studies contributing to this chapter are:

1. To develop a bronchial brushing technique as a model system of lung injury and repair in sheep. Such a technique should allow for reproducible epithelial damage which can be subsequently precisely localized and assessed.
2. To study the histopathological changes of the airways in response to physical injury
3. To investigate the time course of proliferation and repair activities following physical lung injury in different lung segments of the sheep at different time points.
4. To investigate the effect of bronchial brushing on the differentiation status of the airways using a combined AB -PAS technique
5. To develop a simple and reproducible morphometric technique for quantification of structural morphology and proliferation in the airways in response to physical injury

2.3 MATERIALS AND METHODS

2.3.1 Animals

All experimental procedures were subjected to ethical review at the University of Edinburgh and were performed under licence, as specified by the Animals (Scientific Procedures) Act 1986. Nine (n=9) commercially sourced crossbred sheep (Appendix 2.1) aged between 14 and 16 months were used in this study (bodyweight 59.4 ±8.8kg).

2.3.2 Experimental design

Sheep were anaesthetised (Section 2.3.3 in this chapter) in order to facilitate the process of mediating physical injury by application of the bronchial brushing technique. The brushing protocol (Section 2.3.4 in this chapter) was similar to that previously described (Heguy A et al., 2007), consistent, and employed in such a manner as to selectively damage only a small area of epithelium (approximately 15mm x 4mm) located on the ventral aspect of each airway.

Initial discussions in relation to balancing the potential return from individual experiments against the ethical cost of using adequate group sizes necessary to adequately power comparisons between time points after injury, centred on the capability of following changes within the same animals over time. An additional advantage is the facility to study ongoing airway repair responses at several individual spatially disparate sites within the same lung – a refinement that has obvious value in reducing animal numbers. However, the basis of such an approach demands knowledge that changes, in response to separate treatment regimes, occurring at different sites and examined at the same point in time are independent of each other and are largely consistent with changes that would be seen, should all airways be subjected to the same treatment regime. To satisfy this requirement a preliminary study was implemented comparing two protocols that would reflect this comparison.

In Protocol A (n=4), a bronchial brushing was obtained from a single predefined site within the major airways of the anaesthetised sheep, and this procedure repeated on

several subsequent occasions, using separate individual sites each time. As a consequence, at the time of necropsy, sites within each lung could be identified that would presumably reflect the continuum of repair in response to injury.

The assumption in this progressive injury protocol is that different brushing sites are independent of one another and that prior injury in one area of the lung would not impact on the lung as a whole. In order to address the possible influence of the nature of the repair process in individual segments, a second protocol was designed - protocol B.

In Protocol B (n=5), several different sites were brushed within the same lung during one anaesthetic period, and no further lung interventions made until the sheep was killed for necropsy examination. As a consequence several brushing sites at the same stage of healing could be identified within the same lung at necropsy examination - single time point per sheep.

Figure 2.1 depicts the essential features of each protocol. In both protocols, the brushing technique was consistent and employed in such a manner as to damage only a defined portion of each airway. In this manner it was possible to obtain, at post mortem examination, airways at various stages of repair subsequent to physical injury. In protocol A, the repair process was studied at 6 hours, and 1, 3, 7, 14, 21, 28 and 35 days after injury whilst in protocol B, the repair process was studied only from 6 hours up to 14 days post injury.

Details with respect to which sheep were assigned to each protocol and the sampling time points are given in Appendix 2.2a and Appendix 2.2b.

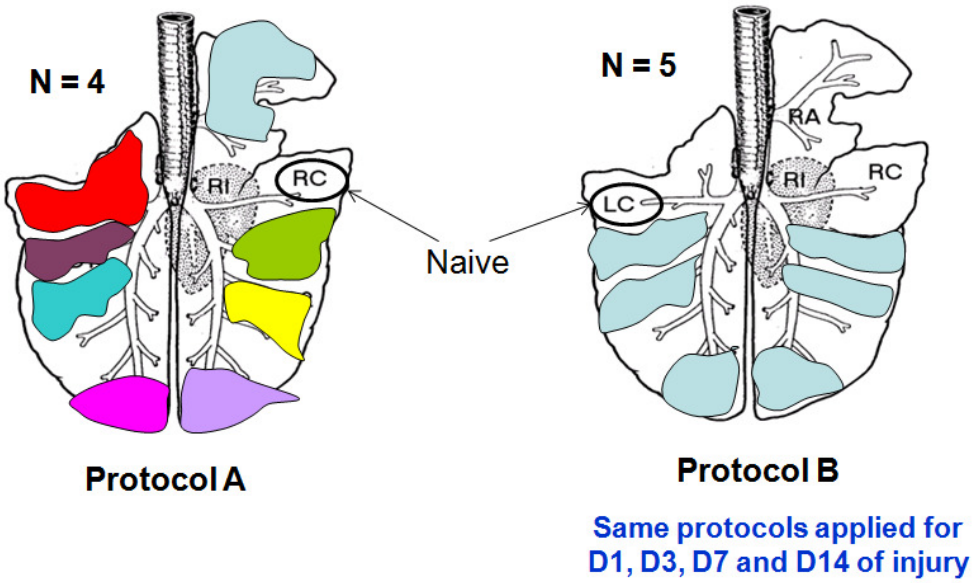


Figure 2.1: Brushing procedures performed according to the protocols
In protocol A, each sheep was anaesthetised on eight separate occasions and on each occasion a separate airway was subjected to bronchial brushing. At the point of necropsy examination airway areas representing the different periods of repair could be identified and sampled. In protocol B, each sheep was anaesthetised on only one occasion with six separate bronchial brushings being performed at six different airways at one single time point. In each sheep for both protocols, samples were also obtained from an unbrushed airway (naïve control). In both protocols, BrdU was administered one hour before the sheep were killed. Different colours in Protocol A indicate the brushing was performed at different time points prior to post mortem whilst in Protocol B, six segments been brushed for each time point.

2.3.3 Anaesthesia

Anaesthesia was induced by intravenous injection of Thiopentone sodium (20mg/kg bodyweight, Thiovet®, Novartis, UK) and thereafter maintained by 2% Isofane in Nitrogen oxide (NO) and oxygen (O_2) (50:50) for the duration of the brushing protocol (15-30 sec/brushing). Sheep were ventilated by positive pressure ventilation (Harvard Apparatus Model 708 (Massachusetts, US)) with tidal volume set at 10ml/kg bodyweight.

2.3.4 Brushing procedure

The bronchoscope was advanced through the endotracheal tube, down the trachea and mainstem bronchi and then into each selected segmental bronchus until the predefined areas selected for induced physical injury was identified. The bronchial brush was then advanced, the sheath retracted and the brush applied to the epithelial mucosal surface. By advancing and retracting the brush in contact with the mucosa, the epithelial surface could be mechanically perturbed by the brush bristles. By changing the position of the tip of the bronchoscope the area of brushing could be precisely controlled. The selected ventral aspect of the airway was injured in this manner. A consequence of repeatedly passing the brush over the mucosal surface was the development of mucosal hyperaemia which served as a visual indication of the area exposed to injury. The brushing procedure was terminated if any noticeable mucosal bleeding occurred.

One hour before the sheep were killed, BrdU was infused via the jugular vein at a dose of 10mg/kg body weight diluted to 10mg/mL in sterile phosphate buffered saline (PBS) solution.

2.3.5 Necropsy

Each sheep was euthanized 1 hour after BrdU infusion by intravenous injection of sodium pentobarbital (Euthatal®, Merial Animal Health Ltd, Essex, UK). Following removal of the heart and lungs the pulmonary circulation was flushed with 3 litres of phosphate-buffered saline (PBS) by connecting a cannula to the pulmonary artery.

The trachea and both right and left main bronchi were incised along their dorsal margins to facilitate identification of the brushed sites on the airway walls of each selected segmental bronchus. The whole circumference of the airway comprising the longitudinal extent of each brushed site was initially carefully dissected free from the surrounding parenchyma. The airway was then separated along the longitudinal plane into two equal trough-shaped proportions, with one portion containing the damaged epithelium along its most ventral aspect and the other opposite portion being naïve to direct injury.

2.3.6 Immunohistochemistry (IHC)

2.3.6.1 Tissues and slides preparation

Both brushed and non-brushed (opposite to the brushed area, and naïve control) sites were carefully identified and dissected away from the surrounding lung tissue. Small intestine and skin samples were also taken as positive controls for BrdU immunostaining. Tissues were fixed with 10% neutral buffered formalin for at least 24 hours before being dehydrated in alcohol and embedded in paraffin. Cut tissue sections (5 µm) were fixed on adhesive slides (X-Tra^{TM}, Surgipath, Peterborough, UK) glass slides, and incubated overnight at 37^{o}C. Tissue sections were then deparaffinized with xylene, rinsed with progressively reducing concentrations of ethanol and immersed in tap water for 5 minutes. Tissue sections were subjected to standard haematoxylin and eosin (HE) and Alcian blue-periodic acid (AB-PAS) histochemical stains, and immunohistochemical staining procedures were directed towards detecting the presence of BrdU and leukocyte common antigen (CD45) (Figure 2.2).

2.3.6.2 IHC for BrdU incorporation

Following deparaffinization, the tissue sections were immersed in 1x PBS for 5 minutes before being acid denatured with 2M hydrochloride acid (HCl) for 8 minutes in a 60^{o}C water bath and rinsed with 1x PBS for 5 minutes. Thereafter the sections were digested with trypsin (pH 7.8) for 30 minutes at 37^{o}C before being washed twice with 1x PBS for 5 minutes and then incubated with pre-diluted normal blocking serum (normal horse serum) for 20 minutes. Subsequently all tissue

sections were incubated overnight with mouse anti-BrdU (1:500 dilutions) antibody at 4°C. To minimize endogenous peroxidase activity, the tissue sections were incubated with 3% hydrogen peroxide (H_2O_2) for 10 minutes after the 24-hour incubation with anti-BrdU. Visualisation protocols were as specified by the manufacturer (RTU Vectastain Universal Elite® ABC commercial kit, Vector Laboratories, Peterborough, UK). Sections were counterstained with haematoxylin, rehydrated and mounted with DPX. Negative control sections, processed and stained as described, were either derived from an animal not exposed to BrdU or from BrdU-treated animals, but with the primary antibody step omitted.

2.3.6.3 IHC for leukocyte common antigen (CD45)

After deparaffinization, the tissue sections were immersed in 1x PBS for 5 minutes before being heated for 30 minutes in a pressure cooker at high-power in an 800W microwave oven. A high pH retrieval solution was used (pH 9.0) [consisting of 1.21g of TRIS base (Sigma, UK), 0.37g of EDTA (Life Technologies, UK), pH adjusted to 9.0, and 0.5ml of Tween-80 (Sigma, UK), made up to 1L with distilled water]. After antigen retrieval treatment, the slides were allowed to cool down for 10 minutes at room temperature before being immersed in distilled water and washed with 1x PBS for 5 minutes. The tissue sections were then incubated with pre-diluted normal blocking serum (normal horse serum) for 20 minutes before being incubated for one hour with mouse anti-ovine CD45 (Serotec, UK) monoclonal antibody (1:500 dilutions) diluted in 1x PBS at 37°C. After the primary antibody incubation, the sections were handled as previously specified for detecting BrdU incorporation, including incubation with 3% hydrogen peroxide, incubation with secondary antibody, incubation with ABC Reagent, diaminobenzidine (DAB) and counterstaining. Negative control slides were prepared as described above, but with the primary CD45 antibody step omitted.

2.3.7 AB-PAS staining

After deparaffinization, the tissue sections were immersed in distilled water for 5 minutes; before being incubated in Alcian blue (Sigma, UK) solution (pH 2.5) for 5 minutes. The sections were then washed for one minute each, in tap and distilled

water respectively, followed by incubation in 1% periodic acid (Molecular Probes, UK) for 10 minutes. The sections were rinsed well for one minute each, in tap and distilled water respectively, and then incubated for 20 minutes in Schiff's reagent (Sigma, UK). The sections were washed in running tap water for 10 minutes, counterstained with haematoxylin for 35 seconds, rinsed with Scott's tap water for 10 seconds, washed with running tap water for 5 minutes, dehydrated with serial degraded alcohol, cleared with xylene and mounted with DPX. Scott's tap water, in acting as a blueing reagent, rapidly and precisely stains the nuclear chromatin and nuclear membrane and as a consequence improves recognition and resolution in histological stains.

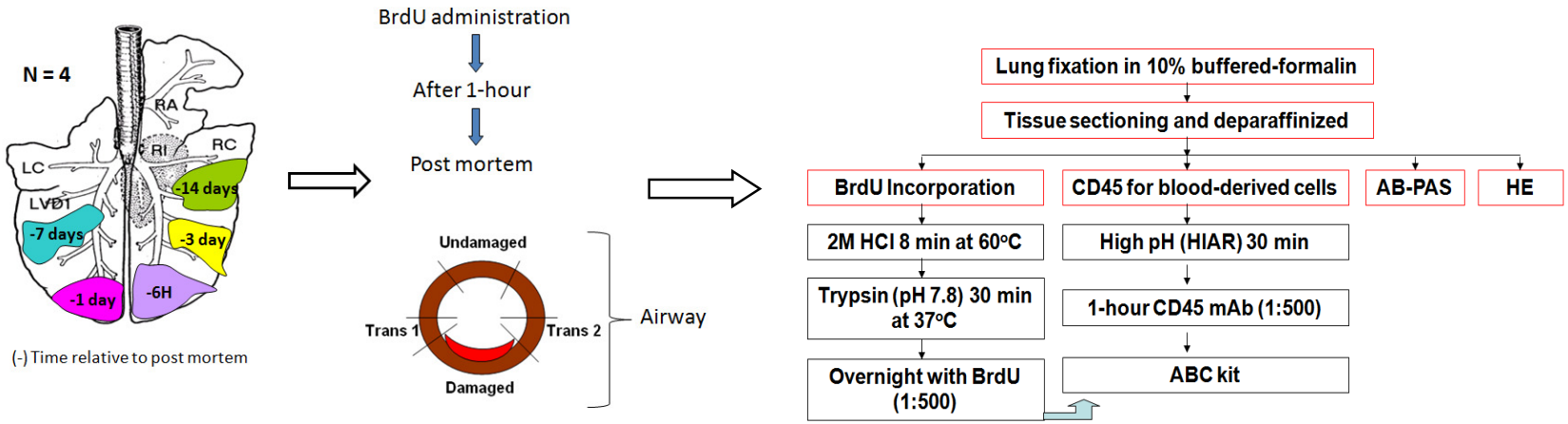


Figure 2.2: A flowchart of the procedures undertaken in this study involving the sheep in Protocol A
(Left - right) The structure of the sheep lung and the segments brushed at different time points i.e up to 14 days of injury (* (-) prior to post mortem). BrdU was administered 1 hour before the sheep was killed and the brushed airway wall areas were identified. The lung segments were carefully mapped and dissected out during the post mortem and fixed with 10% buffered-formalin. After processing, thin (5µM) sections were cut and fixed on coated glass slides. Within each airway cross section prepared in this way, damaged, transitional and undamaged areas were easily identified for further appraisal and quantification. Each section was subjected to IHC for BrdU and CD45 and histochemical appraisal using HE and a combined AB-PAS technique.

2.3.8 Morphometric analysis

In addition to quantifying the extent of cellular proliferative activity using BrdU staining in airway wall sections, preliminary evaluation was directed towards assessing whether airways selected from different lung segments shared similar structural morphology – a necessary prerequisite to interpreting time-dependent changes amongst different lung segments.

The initial analysis of HE-stained sections focussed on the proportion of tissue components in unbrushed sections of airways from the separate sample sites. The components of interest and their definition and locations were as specified by (Anthony B et al., 1994). These components included tissue categories (cartilage, smooth muscle, epithelium, blood vessels) and compartments (mucosal and submucosal regions). Such data collection can be used to ascertain whether, at steady-state, any differences exist between airways from different segments of the lung.

2.3.8.1 Morphometric procedures

The initial aim was to assess the influence of sample site on the structural morphology of the airway wall in different segments of the lung. Subsequent aims involved both a primary appraisal of the nature and extent of any time-dependent changes in cell proliferation in the airway wall and an analysis of the influence of the choice of protocol on the repair response following physical injury. A similar technique for morphometric analysis for HE and BrdU-incorporated cells was applied on the airway sections for both protocols.

In order to facilitate comparisons within and between airway samples, considerable care was taken to identify and dissect the appropriate sections of the relevant airways. Following fixation, these sections were precisely trimmed and arranged within the tissue processing/embedding cassettes in order to maximise the likelihood that airways would be embedded in such a way that tissue sections would consistently be cut in a transverse orientation.

Morphometric measurements for HE-stained sections were performed at a lower magnification of 100X (10x objective) using a multipurpose test grid composed of a combined line and point system (20.4mm square 10 x 10 grid, Olympus UK Ltd, Essex, UK). For estimating the proportion of each of the airway structures and tissues within the airway walls a set of grid points was placed on the section. By counting the number of points falling on each component structure (i.e epithelium, smooth muscle, blood vessel, cartilage, mucosa and submucosal) the volume composition of the structure could be directly estimated.

Whilst BrdU positive staining was generally easily identified at a magnification of 200X (20x objective) as dark-black nuclei on the proliferating cells, where any difficulty was experienced in precisely identifying individual cells increased magnification was utilised to confirm identity. The slides used in the analysis involved only those stained with anti-BrdU antibody. In order to provide a basis for comparison, airways from both treated and untreated (naïve) segments were examined.

Whilst our interest was centred on the area of airway wall directly exposed to injury, further measurements were made on undamaged areas of the airway wall opposite the area of injury within the treated segments and areas lying in the transition regions between these opposite zones. These main areas of interest were thus defined as damaged, transitional and undamaged (Figure 2.2 and Figure 2.3).

Whilst damaged areas were easily identified in the first few days after injury, preliminary studies indicated that, at later time points (days 7 and 14), difficulty could be experienced in reliably identifying these areas at a gross level. However, histologically, by days 7 and 14, those areas previously denuded by the cytology brush were covered by squamous airway epithelial cells, and as a consequence the normal complement of mucus secreting cells was absent. Following physical marking of the slides using lines etched on the surface of the coverslips to delineate the boundaries between zones, the slides were blinded prior to morphometric analysis.

Two compartments of interest, mucosa and submucosa, within the airway wall were defined according to the aforementioned criteria proposed by Anthony B et al., (1994). Thus, the mucosal region consisted of the epithelium, basement membrane, and lamina propria, whilst the submucosal region began at the outer border of the lamina propria and included the smooth muscle layer (Anthony B et al., 1994).

In order to quantify the proportion of different tissue categories within the airway wall, a point counting method was used. In this regard, the principle of point counting relies on counting the number of points on a multipurpose test grid graticule that overlies the area of interest, and relating these points to the total number of points in the grid to gain an approximation of the proportion of the area of interest within the grid area. The volume densities of tissue categories and the compartments of interest (mucosal and submucosal) were calculated by relating the number of points landing on each tissue category and compartment to the total number of points counted.

By contiguous placement of consecutive grids, the whole of the airway in question or subordinate areas of interest were subject to point counting. For the purpose of consistency and reproducibility, the counting began with the grid placed on the marginal edge of the damaged area and, following counting, this was moved to the next contiguous grid area within the damaged area. A similar approach was used for both transitional zones flanking the damaged area and the undamaged area opposite (Figure 2.3). The point counting data were recorded in written form and transcribed into a spreadsheet for analysis.

It may be appreciated that the number of fields studied in each section was potentially determined by several factors, but notably the area of initial damage induced as a consequence of the brushing procedure and the time subsequent to injury. In the latter regard, it was anticipated that the area of apparent damage would decrease relative to its initial dimensions, and as a consequence the variance associated with data pertaining to later time points would be greater. In point of fact, as a consequence of the sizes of the airways in question and the magnification used,

the number of fields measured ranged from 2 - 4. All measurements were conducted with the slides independently blinded prior to examination.

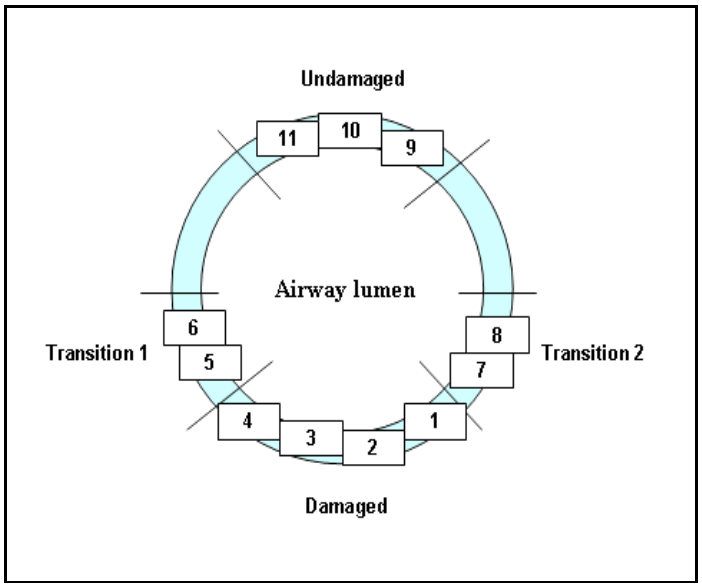


Figure 2.3: Diagram showing the field selection procedure undertaken for morphometric analysis
For the purpose of consistency and reproducibility, the counting began with the grid placed on the marginal edge of the damaged area (number 1) and, following counting, the square grid was moved to the next contiguous area within the damaged area (numbers 2, 3 and 4). In this manner the whole area of damage was subject to evaluation. The same procedures were applied in relation to the flanking transitional areas, and the undamaged area opposite the damaged airway.

2.3.8.2 Statistical analysis

The experimental design, involving repeated measurements on various tissue samples from different segments within the same sheep, facilitated the application of a mixed effect model analysis similar to that described by Paterson S and Lello J, (2003). The individual samples were derived from different segments, each relating to different time points. Material collected from sheep involved in both protocols were subjected to the morphometric analysis to measure the effect of the segment on the distribution of the airway tissues (HE-stained sections) whilst only data collected from sheep involved in protocol A were analysed to assess the extent and distribution of the proliferative cell response following injury (BrdU-stained sections). Since only one sheep was assigned for each time point, therefore no further statistical analysis was carried out.

The model was fitted using a generalized linear mixed effect model, using penalized quasi-likelihood (PQL) and, binomial (HE proportion data) and Poisson (BrdU count data) errors. In these analyses (HE and BrdU), the sheep's ID was used as a random effect in order to control for any pseudoreplication and to minimise the noise associated with variation between observations and/or measurements on the same sheep, thereby allowing for the significance of factors such as segment identity and sheep to be tested within the fixed component of the model (Crawley, 2007, Paterson S and Lello J, 2003). In these models, the total number of grids landing on the areas measured (i.e mucosal, submucosal, damaged, transitional and undamaged) was used as a covariate in order to control/reduce variation associated with the measuring process such as the size of the airways measured or the number of grids placed on the measured areas. The data was analysed using the freely available statistical software environment R (The R Foundation, Software version 2.9.0, Vienna, Austria). In all cases, $p<0.01$ for HE and $p<0.05$ for BrdU counting were considered statistically significant.

2.4 RESULTS

2.4.1 Histopathology of the airways after injury: HE staining

2.4.1.1 Quantitative morphometric analysis: Is airway structure consistent between segments?

In assessing the similarity of structural morphology of different segments of the sheep airway, the proportions of each tissue categories were plotted as box plots [(Figure 2.4) and (Figure 2.5)]. Overall, it is clear that in this model, with the Sheep ID as a random effect and segments RA (protocol A) and RVD1 (protocol B) as reference levels, there was no significant difference ($p>0.01$) in the proportion of each separate tissue component between segments of the lung in these groups. However, in Protocol A sheep, for the vast majority of cases there were no differences between any of the 8 segments for the 6 tissue categories ($p>0.017$), except for RCD (mean=18.9%) and LC (33.9%) in mucosa ($p=0.002$) and LVD1 (9.1%) and RA (mean=3.5%) in blood vessels ($p=0.005$) (Figure 2.4). The R outputs are provided in the Supplement 1 on the CD. This analysis also revealed that the size

of the mucosal and submucosal compartments did not differ between segments. In order to test the validity of the reference level, in protocol A the analysis was repeated in order to test the proportion of smooth muscle between different segments by choosing LCD or RCD as reference levels. The results revealed findings essentially to the same as those obtained when using RA as a reference level.

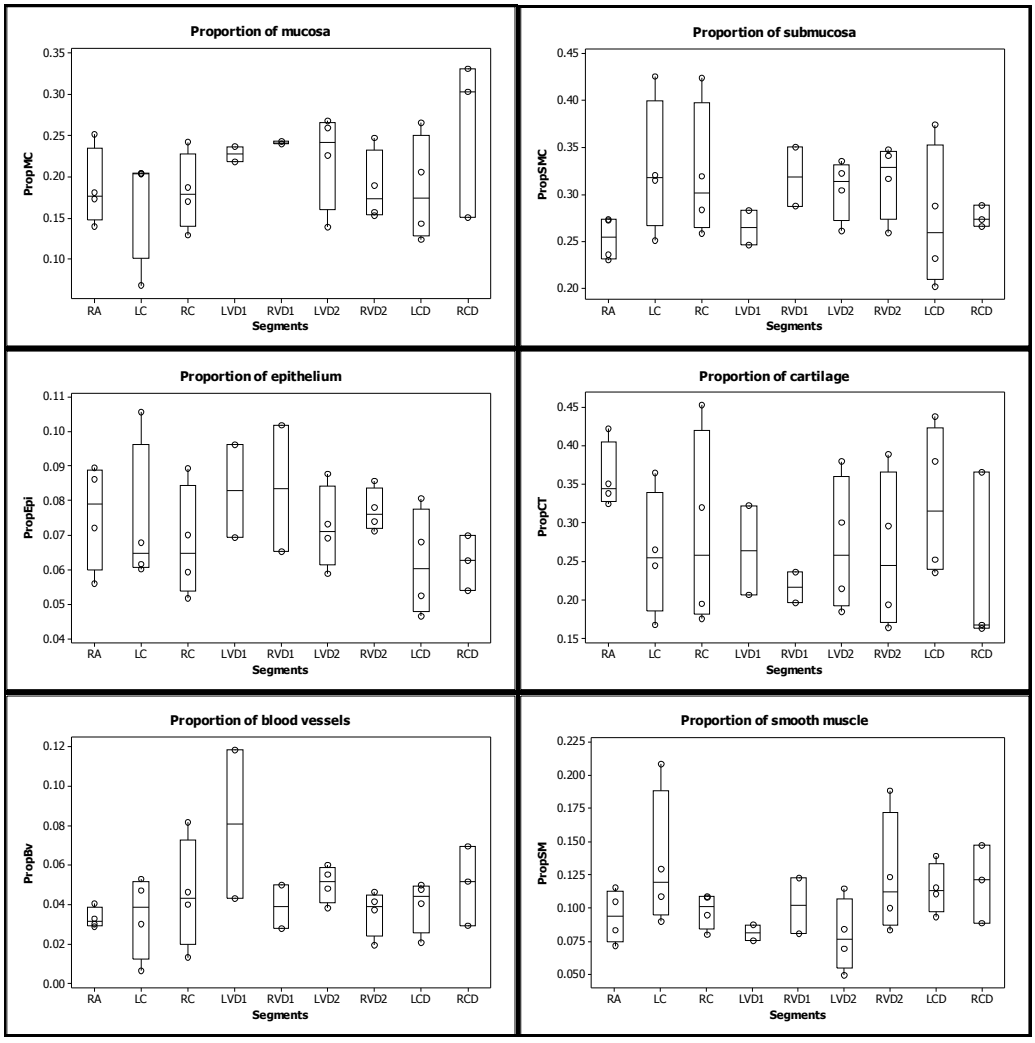


Figure 2.4: The box plots for the proportions of the tissue categories between segments of the lung sections for the group A sheep
The statistical analysis was carried out using a generalized linear mixed model (family binomial). The results revealed that there was no significant difference in the relative proportions of the tissue categories between different segments of the lung (using segment RA as the reference level). A $p<0.01$ level was considered statistically significant. Symbols represent proportion data for the number of points landing on tissues for each area measured relative to the total number of grid points landing on each tissue measured. Upper and lower box plot margins represent the interquartile range; middle bar indicates the median. The points outside the ends of the whiskers are outliers.

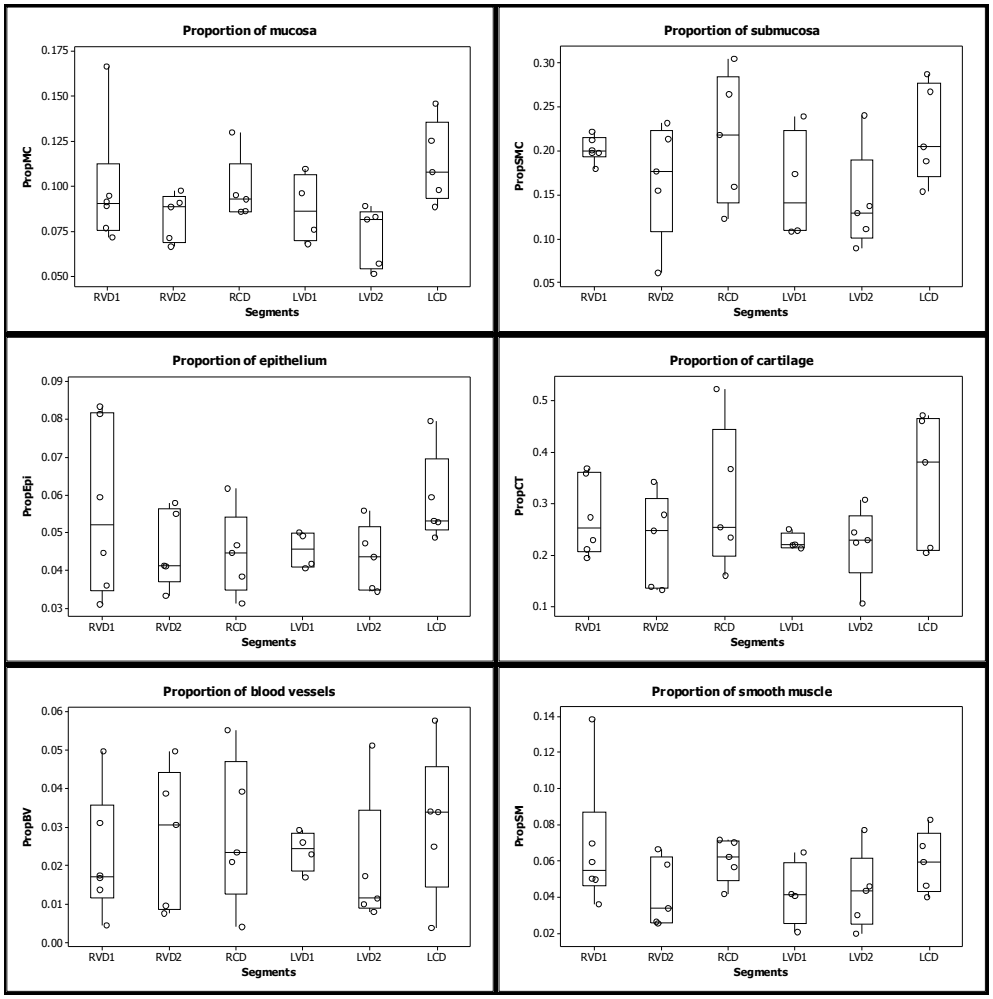


Figure 2.5: The box plots for the proportions of the tissue categories between segments of the lung sections for the group B sheep

The results revealed that there was no significant different of the tissues of interest between difference segments of the lung as compared to the segment of the RVD1 as a reference level. A $p<0.01$ was considered statistically significant. Symbols represent proportion data for the number of points landing on tissues for each area measured relative to the total number of grid points landing on each tissue measured. Upper and lower box plot margins represent the interquartile range; middle bar indicates the median. The points outside the ends of the whiskers are outliers.

2.4.1.2 Histopathological assessment of airway damage and repair

Whilst in most instances the brushed sites were easily identifiable during gross necropsy examination, in some instances, particularly at later time points, it became more difficult to visually distinguish the sites of injury. However, the standardised brushing protocol and careful mapping of brushed sites facilitated accurate sampling of the appropriate areas. At a histological level, the sites of brushing were readily apparent, as were time-dependent changes in histopathological appearance. The pathophysiological changes of the airways following a bronchial brushing were compared to the normal morphology of the airway (naïve and undamaged sites (Figure 2.6)).

At six hours after injury (Figure 2.7) mucosal loss and epithelial attenuation (Figure 2.7; left panel) were apparent together with submucosal oedema (Figure 2.7; bottom right panel). There were small numbers of neutrophils, lymphocytes and plasma cells percolating around blood vessels and submucosal glands at the site of injury (Figure 2.7; top right panel).

Twenty-four hours after injury, within the brushed sites, there was evidence of respiratory mucosal loss (Figure 2.8; top left panel) with a superimposed layer of fibrin, neutrophils and erythrocytes. At the margins of the brushed sites the respiratory epithelium was attenuated, ranging from cuboidal to squamous (Figure 2.8; top panels). In some areas only a single layer of basal epithelial cells was apparent (not depicted). There was oedema in the deep submucosa and mucosa, characterised by increased spacing and clearing between collagen bundles in the deep mucosa and also between the inflammatory cells in the mucosa (Figure 2.8; bottom left panel). Inflammatory cells consisted of small numbers of neutrophils, lymphocytes and plasma cells (Figure 2.8; bottom right panel). In some areas there were also small numbers of eosinophils (not depicted). One or two submucosal glands were surrounded by small numbers of neutrophils (Figure 2.8; bottom right panel) at the site of injury.

Three days post injury (Figure 2.9) the mucosa at the brushed site was segmentally lost (ulcerated) with a superimposed layer of fibrin, blood and neutrophils (Figure 2.9; top panels). The ulceration extended to the superficial aspect of some submucosal glands, many of which were lined by attenuated epithelial cells and contained a small number of neutrophils (Figure 2.9; bottom left panel). The submucosa and the lamina propria were infiltrated by small to moderate numbers of neutrophils, eosinophils, lymphocytes and plasma cells (not depicted) and there was evidence of mild neovascularisation and fibroplasia within the brushed sites (Figure 2.9; bottom right panel). At the margins of the brushed site, there was mild epithelial hyperplasia with neutrophilic exocytosis (Figure 2.9; bottom right panel).

Seven days post injury (Figure 2.10), the epithelium at the brushing site consisted of stratified squamous epithelial cells consistent with squamous metaplasia (Figure 2.10; top right panel). There was neutrophilic exocytosis across the surface of the epithelium and the mucosa was thickened by granulation tissue (Figure 2.10; bottom panels), which replaced the normal smooth muscle bundles in this region (Figure 2.10; left panels). This tissue was infiltrated with lymphocytes and plasma cells (not depicted). Squamous epithelium lining some submucosal glands was also identified (not depicted in the figure).

Fourteen days (Figure 2.11) after injury, in some areas of the epithelium over the brushed areas (Figure 2.11; left panel), there was a slight reduction in the height of the mucosal epithelial cells with a mild reduction in goblet cell number relative to the unbrushed sites (unbrushed site is not depicted in the figure). Within the lamina propria there were small areas of fibrosis present (Figure 2.11; right panel). Mild attenuation of submucosal gland lining epithelial cells also remained.

Sections prepared from brushed areas from 21 and 35 days post injury indicated that complete repair had occurred (data not shown).

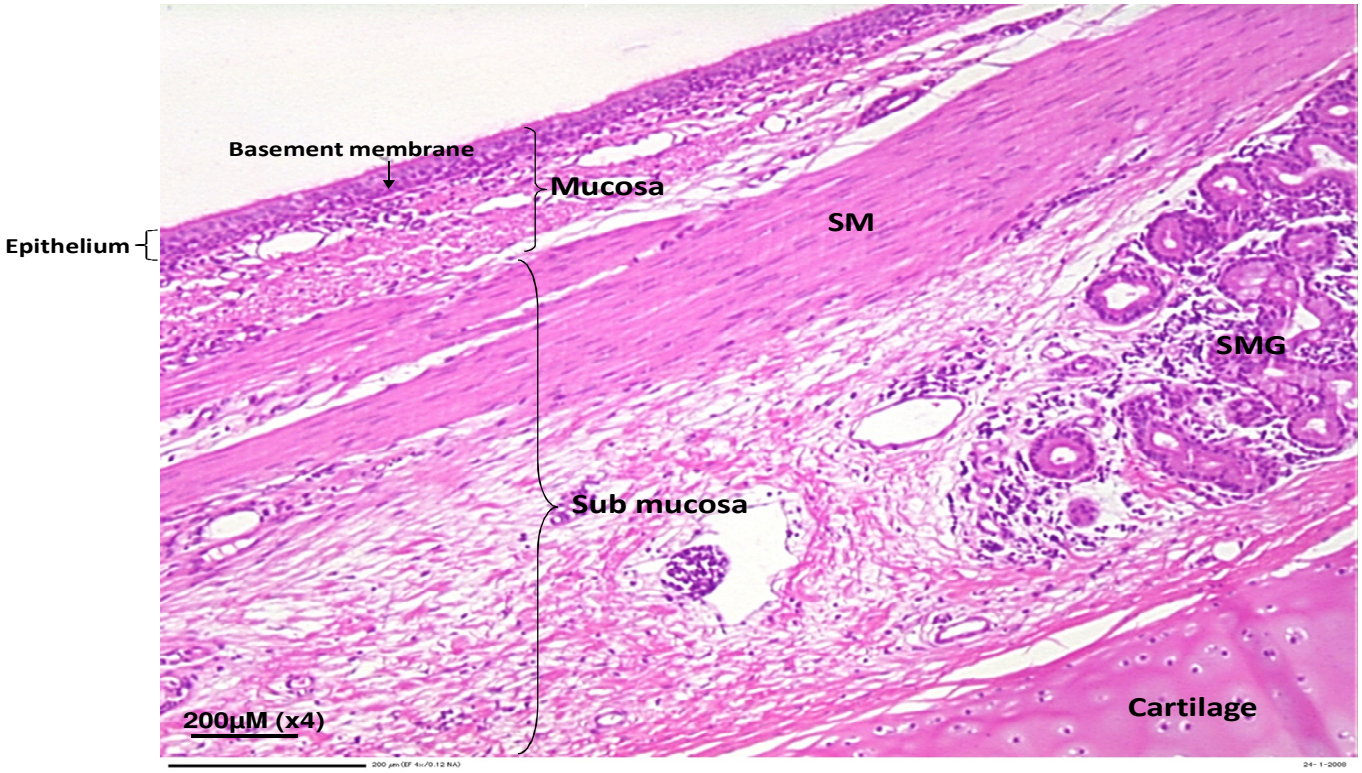


Figure 2.6: Photomicrograph depicting histology typical of the normal airway structure
A cross-section of the normal airway structure stained with HE. The airway consists of mucosal and submucosal compartments as described by (Anthony B et al., 1994) (SM = smooth muscle; SMG = submucosal gland)

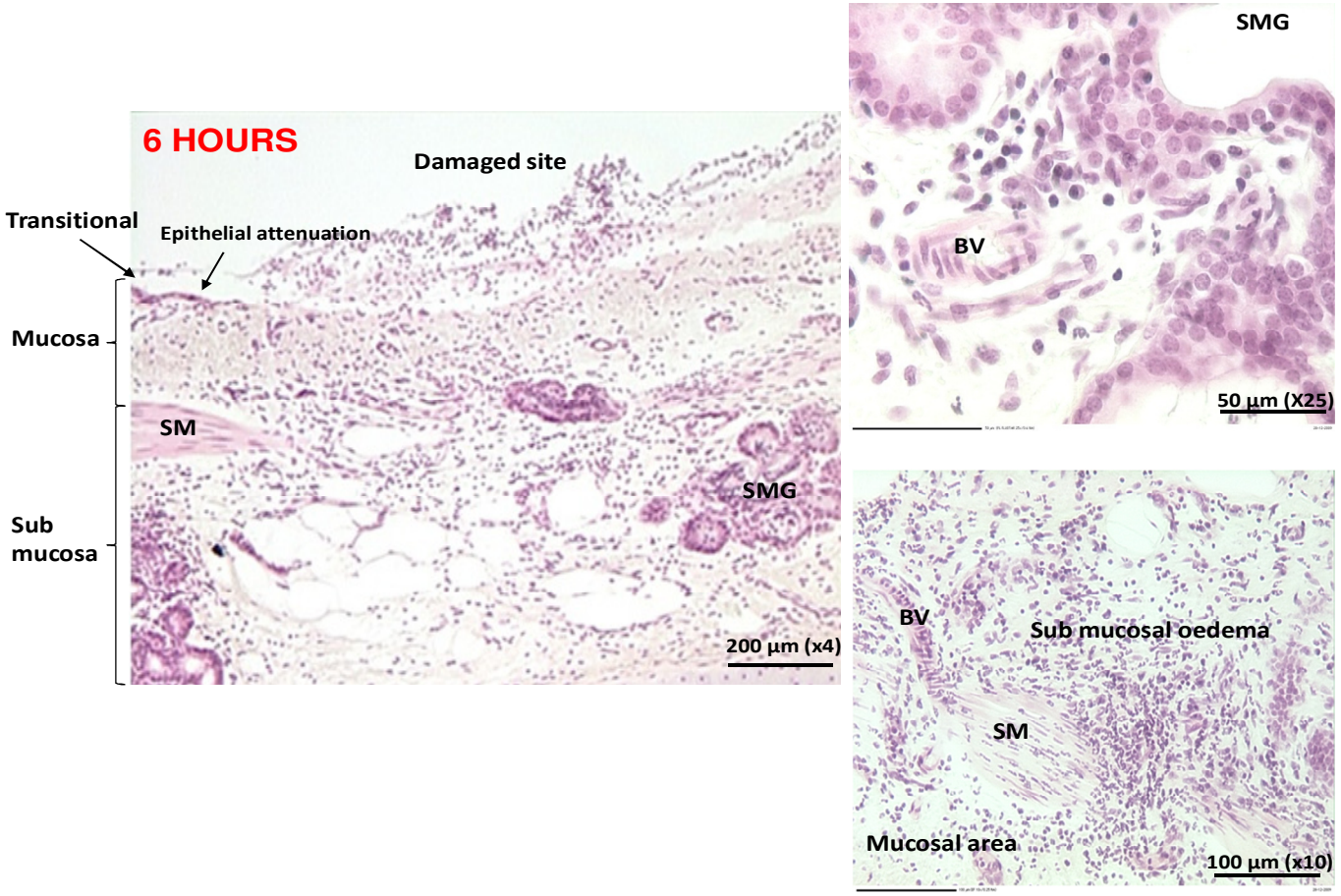

6 HOURS
Damaged site
Transitional
Epithelial attenuation
Mucosa
SM
Sub mucosa
SMG
200 µm (x4)
SMG
BV
50 µm (X25)
BV
Sub mucosal oedema
SM
Mucosal area
100 µm (x10)

Figure 2.7

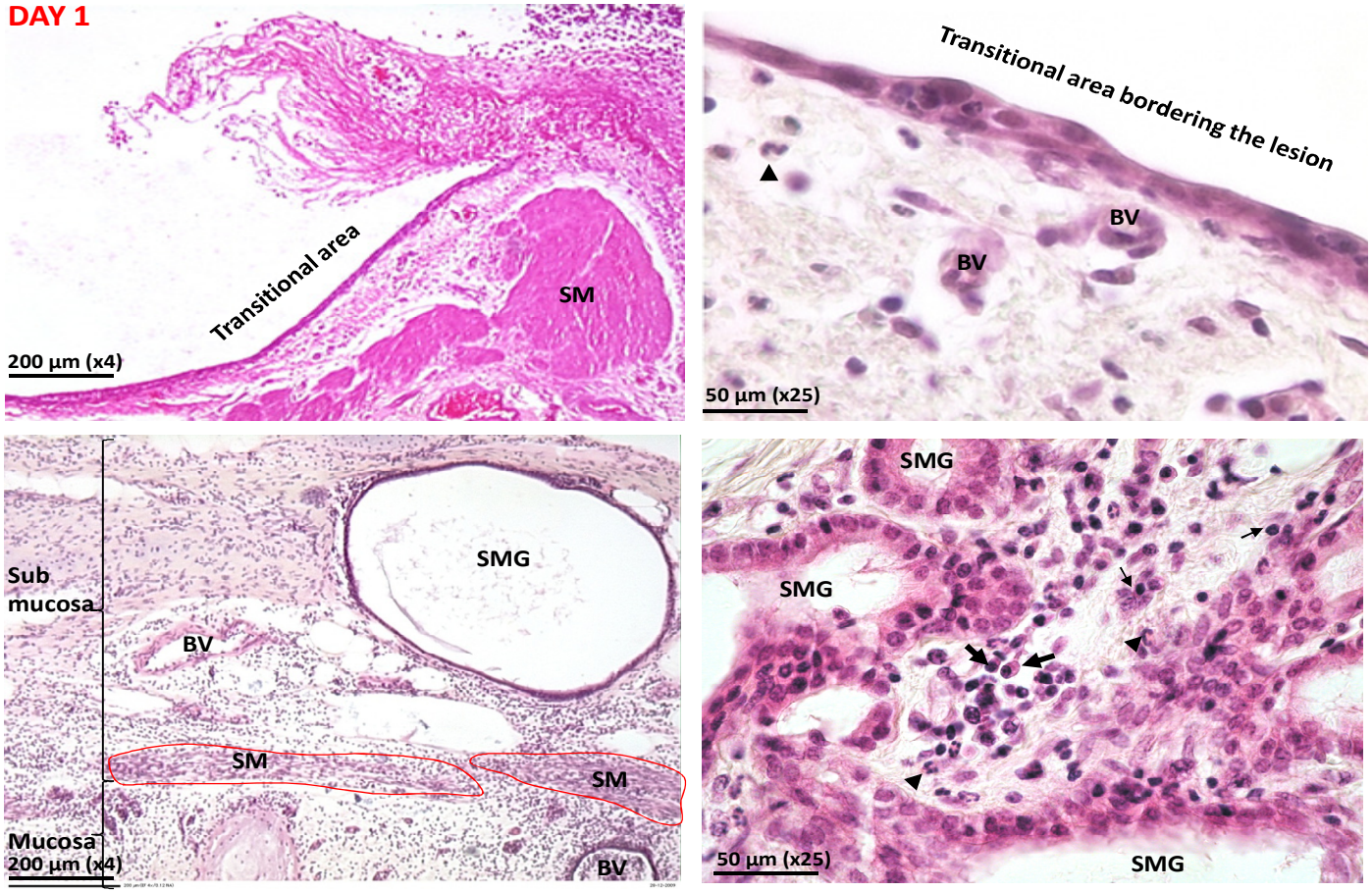

DAY 1
Transitional area
SM
200 µm (x4)
Transitional area bordering the lesion
BV
BV
50 µm (x25)
Sub
mucosa
SMG
BV
SM
SM
Mucosa
200 µm (x4)
BV
SMG
SMG
50 µm (x25)
SMG

Figure 2.8

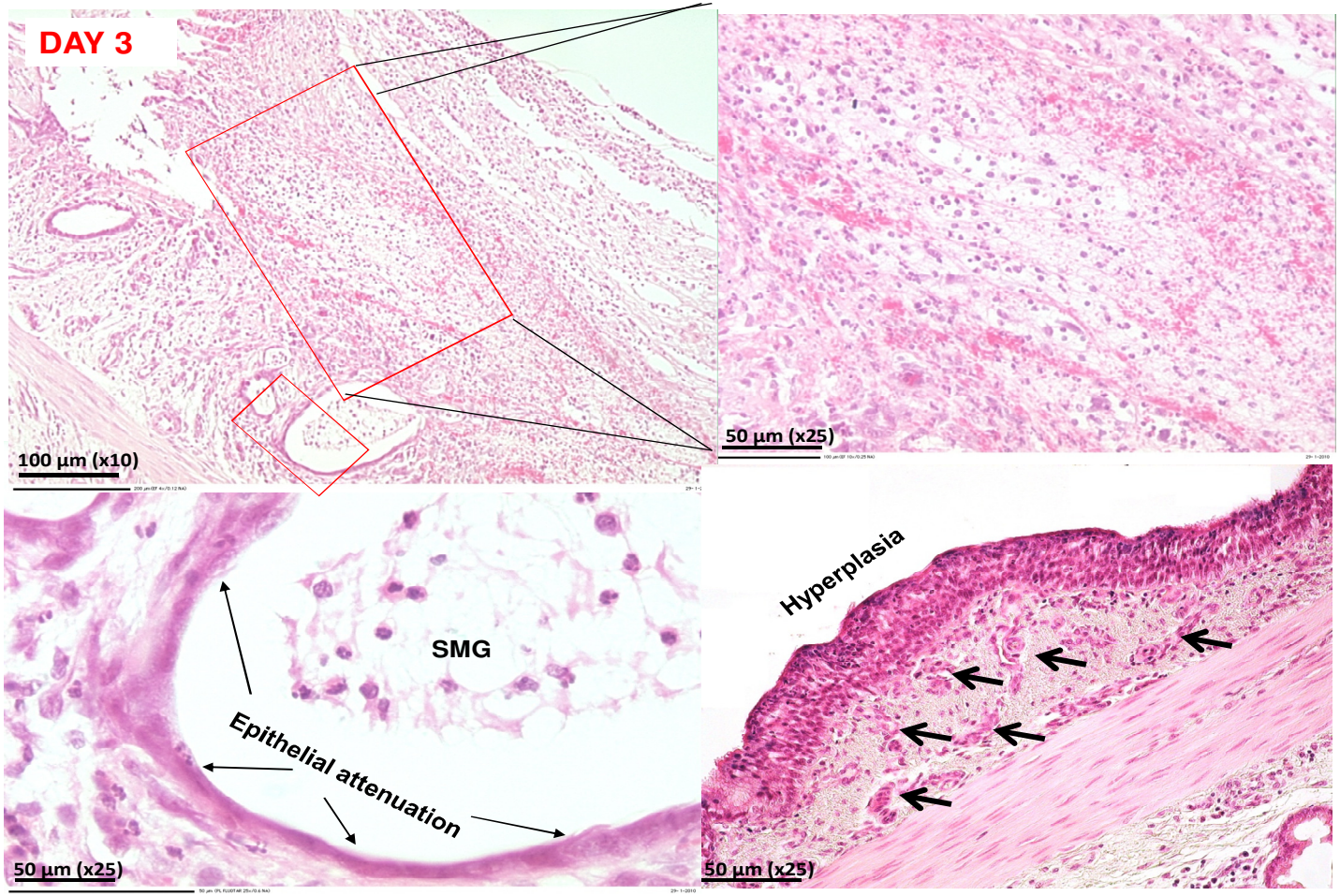

DAY 3
100 µm (x10)
50 µm (x25)
SMG
Epithelial attenuation
50 µm (x25)
Hyperplasia
50 µm (x25)

Figure 2.9

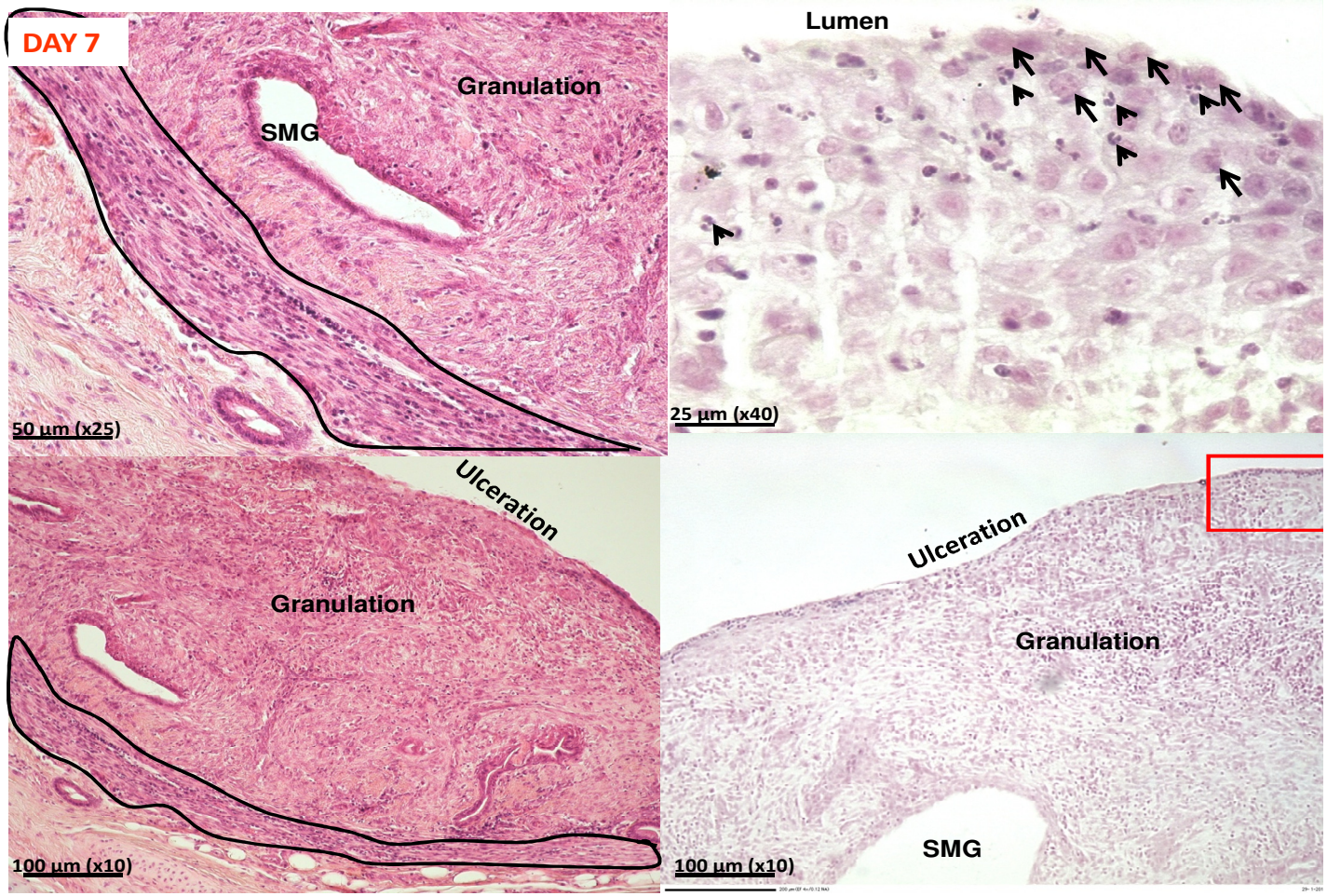

DAY 7
Granulation
SMG
50 µm (x25)
Lumen
25 µm (x40)
Ulceration
Granulation
100 µm (x10)
Ulceration
Granulation
SMG
100 µm (x10)

Figure 2.10

DAY 14

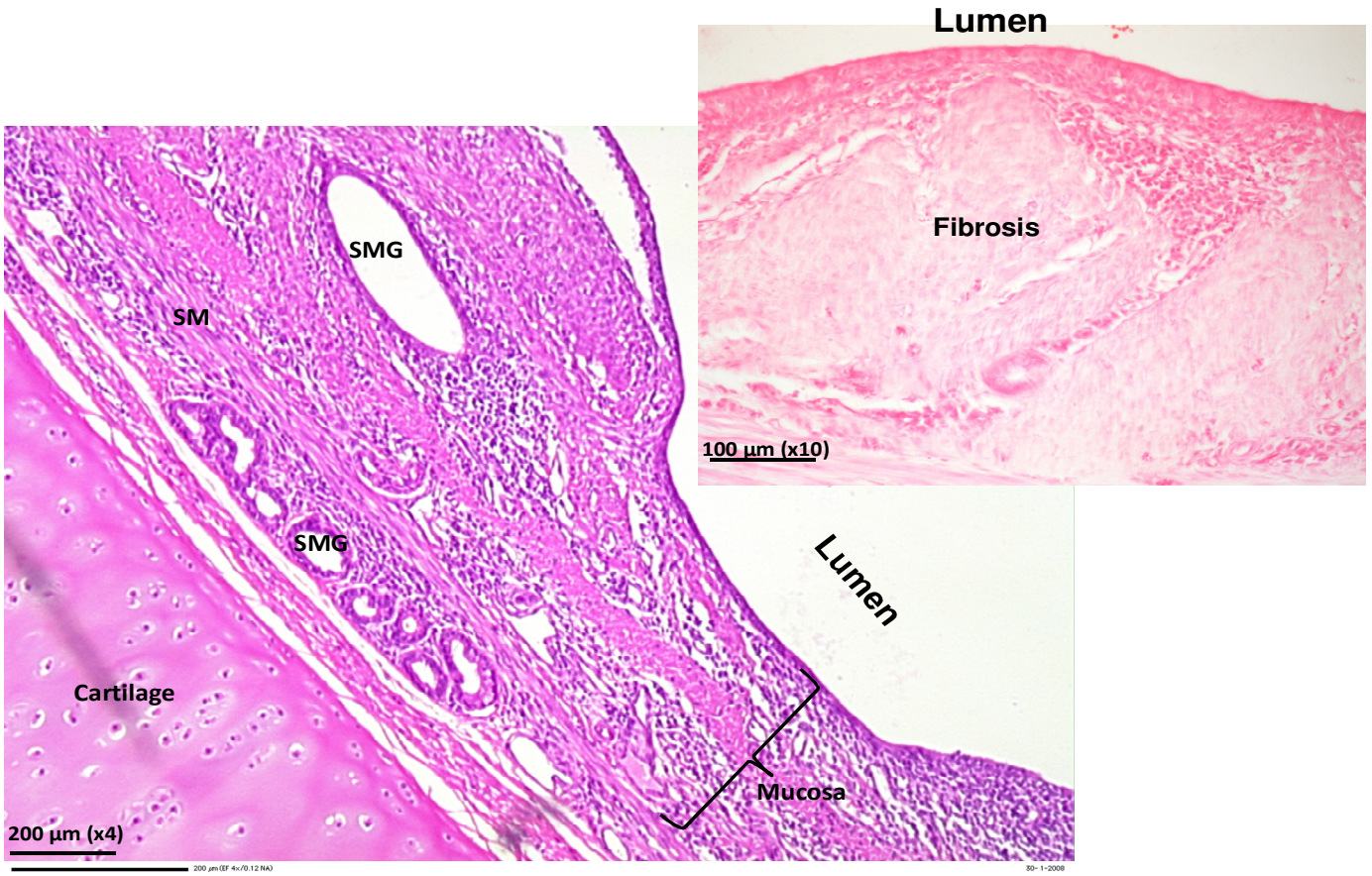


Figure 2.11

Figure 2.7: A Photomicrograph depicting histopathology typical of the repair process underway 6 hours after injury (HE staining)
At six-hour post injury, mucosal loss and epithelial attenuation are apparent, together with submucosal oedema (bottom right panel). There is a simple layer of cells remaining at the transitional area, interpreted to be dedifferentiated epithelial cells (left panel). Small numbers of neutrophils (arrow heads), plasma cells and lymphocytes infiltrate around blood vessels and submucosal glands at the site of injury (top right panel). (SM= smooth muscle; SMG=submucosal gland; BV=blood vessels)

Figure 2.8: Photomicrograph depicting histopathology typical of the repair process underway one day after injury (HE staining)
Twenty-four hours after injury, there is loss of the respiratory mucosa with a superimposed layer of fibrin, neutrophils and blood confined to the brushed areas (top left panel). At the margins of the site the respiratory mucosa is attenuated from cuboidal to squamous (top right panel) and, in some areas, only a lining of basal cells is apparent. The mucosa and superficial submucosa are thickened by oedema and infiltrated by moderate numbers of neutrophils and eosinophils (bottom left panel) which also tended to surround and infiltrate the lumina of submucosal glands (bottom right panel). Arrows show the neutrophilic infiltration. Arrows head show the neutrophils, thick arrows show plasma cells and thin arrows show lymphocytes. (SM=smooth muscle; SMG=submucosal gland; and BV=blood vessels). The marked areas show regenerating smooth muscle tissue.

Figure 2.9: Photomicrograph depicting histopathology typical of the repair process underway three days after injury (HE staining)
By day 3 of injury, the brushed site is covered by a thick layer of fibrin, blood and neutrophils (top panels) which extends to involve the superficial aspects of the submucosal glands, many of which are lined with attenuated epithelial cells and contain a small number of neutrophils (bottom left panel). Within the mucosa there is evidence of neovascularisation (bottom right panel, arrows) and mild fibroplasia. The intact mucosa is attenuated but there are also areas which are thickened, due to hyperplasia (bottom right panel). The large rectangle in the top left panel is shown at higher magnification in the top right panel, whilst the smaller rectangle is shown at higher magnification in the bottom left panel. (SMG=submucosal gland)

Figure 2.10: Photomicrograph depicting histopathology typical of the repair process underway seven days after injury (HE staining)
Seven days post injury, the epithelium at the brushing site consists of stratified squamous epithelium (top right panels; arrows) in which there is also neutrophilic exocytosis (top right panel; arrow heads). The submucosa is focally thickened by granulation tissue (bottom panels) which, in the submucosal area, appeared to displace and/or replace the normal smooth muscle bundles in this region (left panels) (compare this to the normal smooth muscle morphology in Figure 2.6). The areas outlined on the left panel indicate smooth muscle tissue. The small square area outlined in the bottom right panel is shown at higher magnification in the upper right panel. This area shows epithelial squamous metaplasia (SMG=submucosal gland)

Figure 2.11: Photomicrograph depicting histopathology typical of the repair process underway 14 days after injury (HE staining)
Figure shows, the damaged site fully covered with an epithelial layer at fourteen days after injury (left panel) although some segmental abnormalities in the morphology remain. On occasion the mucosa demonstrated small areas of fibrosis (right panel) with some persistent evidence of submucosal gland epithelial attenuation (not depicted). (SMG=submucosal gland; SM=smooth muscle)

2.4.2 Analysis of immunohistochemistry (IHC)

2.4.2.1 BrdU incorporation

Skin and intestinal tissue samples were examined as positive controls for BrdU immunohistochemistry. Cells positive for BrdU were easily identified in sections derived from these samples (Figure 2.12). In the small intestine, the large numbers of positive cells were confined mainly to the crypts while in skin sections, the positive cells were observed in both follicle bulbs and skin epithelium. Negative control slides demonstrated no positive signal in skin, gut or airway sections. For the sheep not been exposed to BrdU, there was no BrdU antigen present in the sheep airway tissue sections. When the slides were stained with BrdU antibody, no positive staining was observed in which the IHC staining for this slide was similar to that seen in the airway tissue sections with BrdU antibody omitted. Both negative control slides are shown (Figure 2.13).

In relation to the airway samples, in both protocols the specific location and number of proliferating cells varied with time of injury. There was little evidence of BrdU positivity in the undamaged region (Figure 2.14, arrows). In the brushed airway sections, BrdU-labelled cells were identified as early as six hours post-injury; however these BrdU-labelled cells were few in number and sparsely scattered through the mucosa and submucosa, not only in the damaged areas but also in transitional and undamaged areas (Figure 2.15, arrows). Figure 2.16 shows the distribution of BrdU-positive cells at days 1, 3 and 7 post injury. At day 1, BrdU-labelled cells were visible within the airway epithelium in clusters flanking either side of the injury ie in the transitional regions. BrdU positive cells were also found underneath the damaged epithelium in the mucosa and submucosa – particularly in the submucosal glands (SMGs). Day 3 distribution patterns of BrdU-labelled cells were similar to those seen on day 1 with respect to epithelial and SMG localisation (not depicted) but in addition large numbers of labelled cells, including those with morphology consistent with fibroblasts and endothelial cells were present in the underlying mucosa at this time point. By day 7 fewer BrdU labelled cells were observed than day 3. In particular very few BrdU-labelled cells were present in the squamous epithelium formed over the lesion. BrdU-labelled cells were still found at

day 14 especially in the sub-epithelial region (Figure 2.17). At later time points (days 21 to 35 post injury), there was very little evidence of BrdU uptake in the sections (data not shown).

In order to compare the distribution pattern of proliferative cells between protocols (Protocols A and B), the total number of BrdU-positive cells were plotted for both protocols and according to the areas of interest and the time relative to the point of brushing injury (Figure 2.18). Although the distribution pattern of BrdU-positive cells shared broad similarities between protocols one noticeable quantitative difference between these two protocols was that the number of BrdU-positive cells peaked at day 3 post injury in the protocol A sheep whilst it peaked earlier (day 1) in the protocol B sheep especially at the damaged areas (Figure 2.18). Although there was evidence of a proliferative cell response to injury in the undamaged areas for both protocols it was noticeably lower for protocol B than for protocol A. When analysed according to time point, compartments (mucosa and submucosa) or area of interest (damaged, undamaged and transitional), in the protocol A sheep (Figure 2.19), there were significant differences in the numbers of BrdU-positive cells observed between compartments (submucosa (geometric means=18.0270) less than mucosa (geometric means=33.401), p=0.000) – (Figure 2.19; left panel), between areas of interest (damaged (geometric means=53.637) and transitional (geometric means=31.441) greater than undamaged (geometric means=8.536), p=0.001 for both –(Figure 2.19); middle panel)) and between time points (day 3 (geometric means=55.013) greater than naïve (geometric means=9.277), p=0.008) – (Figure 2.19; right panel). The details of the R outputs for BrdU data analyses are in the Supplement 2 on CD.

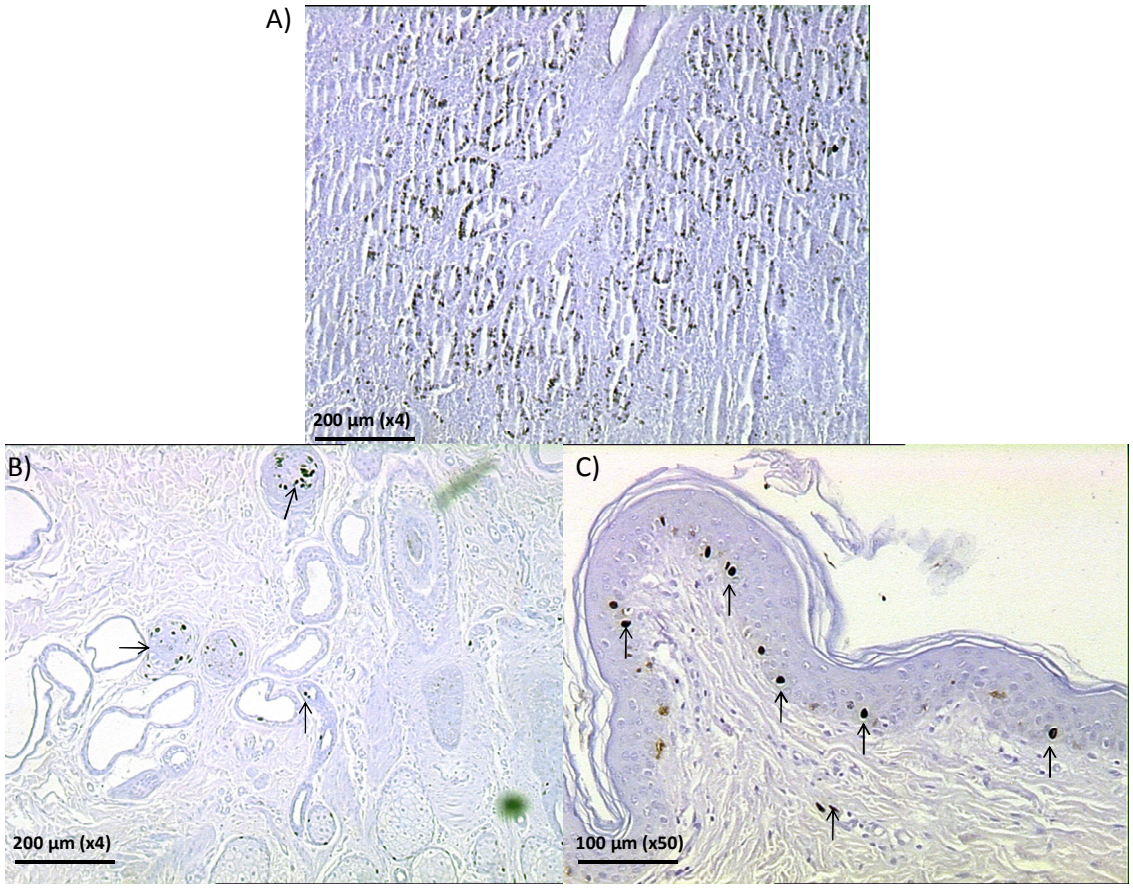


Figure 2.12: A cross section of intestinal and skin tissues stained with BrdU

Cells positive for BrdU were easily identified in sections derived from intestinal and skin samples. In the small intestine (top panel), the positive cells were confined mainly to the crypts while in skin sections, the positive cells were observed in both follicle bulbs (bottom left panel) and skin epithelium (bottom right panel). Arrows show the BrdU-labelled cells.

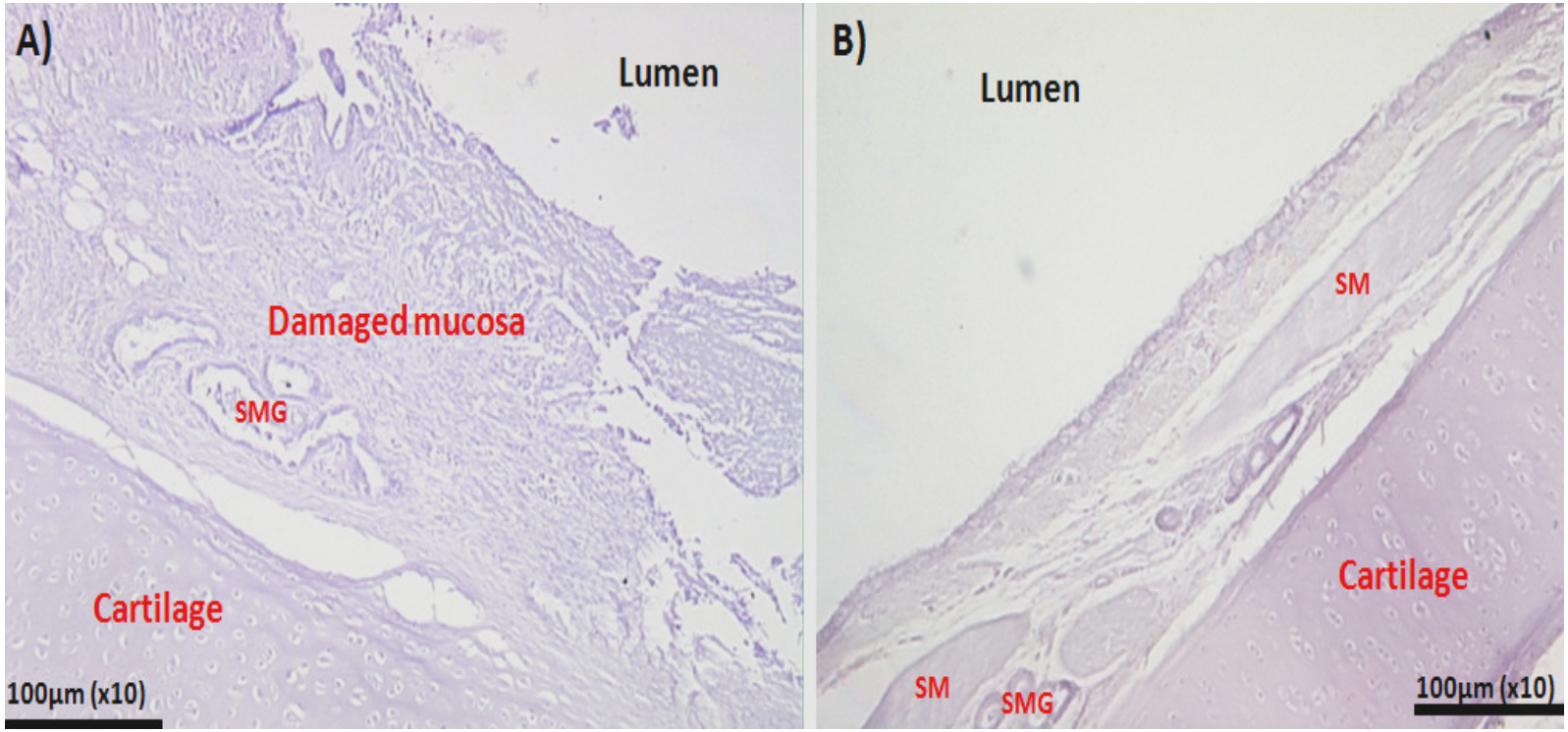


Figure 2.13: Negative control slides
In this instance the sections were derived from an animal exposed to BrdU but the primary antibody step was omitted from the staining process (panel A) and from sheep which haven't been exposed to BrdU but stained with BrdU antibody (1:500 dilution) during the staining (panel B). In both slides, no BrdU-positive cells were observed.

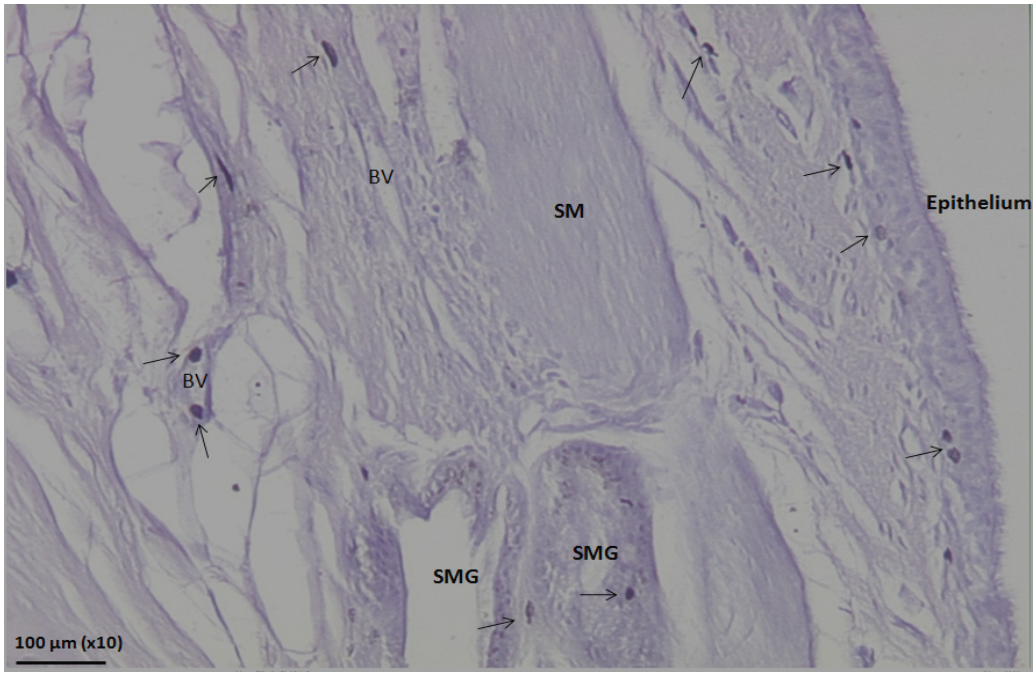


Figure 2.14: A naive airway tissue section stained with BrdU
Few BrdU-positive cells were observed in undamaged airway tissue sections. These cells were sparsely scattered at the basal area of the epithelium, and in the proximity of submucosal glands and blood vessels in both mucosal and submucosal compartments. Arrows show the BrdU-labelled cells. (SMG=submucosal gland; SM=smooth muscle; BV=blood vessels)

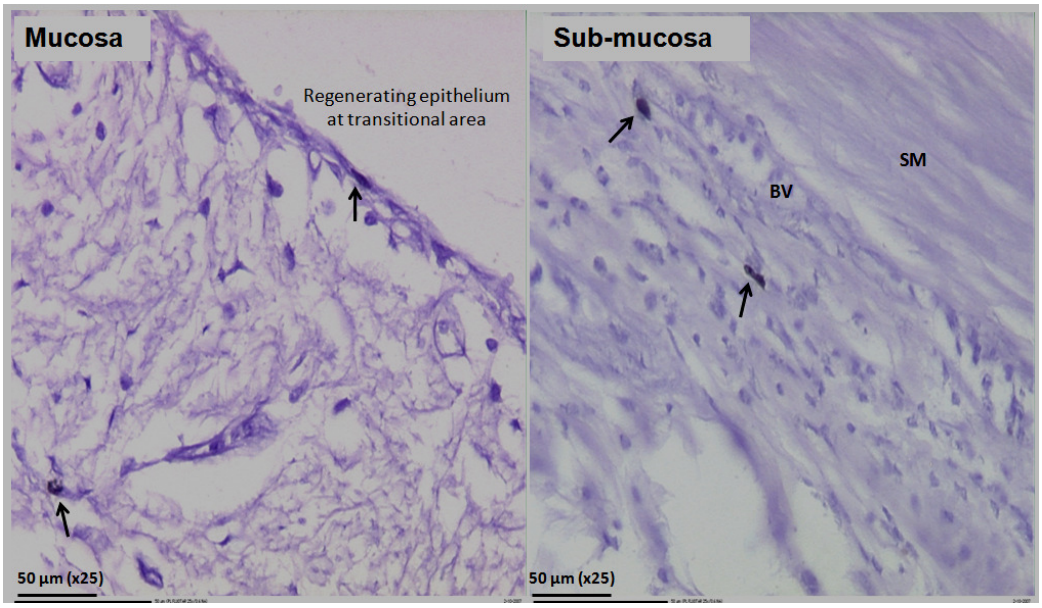


Figure 2.15: The airway tissue sections at 6 hours post injury stained with BrdU
BrdU-labelled cells were sparsely scattered both in the epithelium bordering the lesion and in the proximity of blood vessels (mucosal region – left panel, and submucosal region – right panel. Arrows show the BrdU-labelled cells. (SM=smooth muscle; BV=blood vessels)

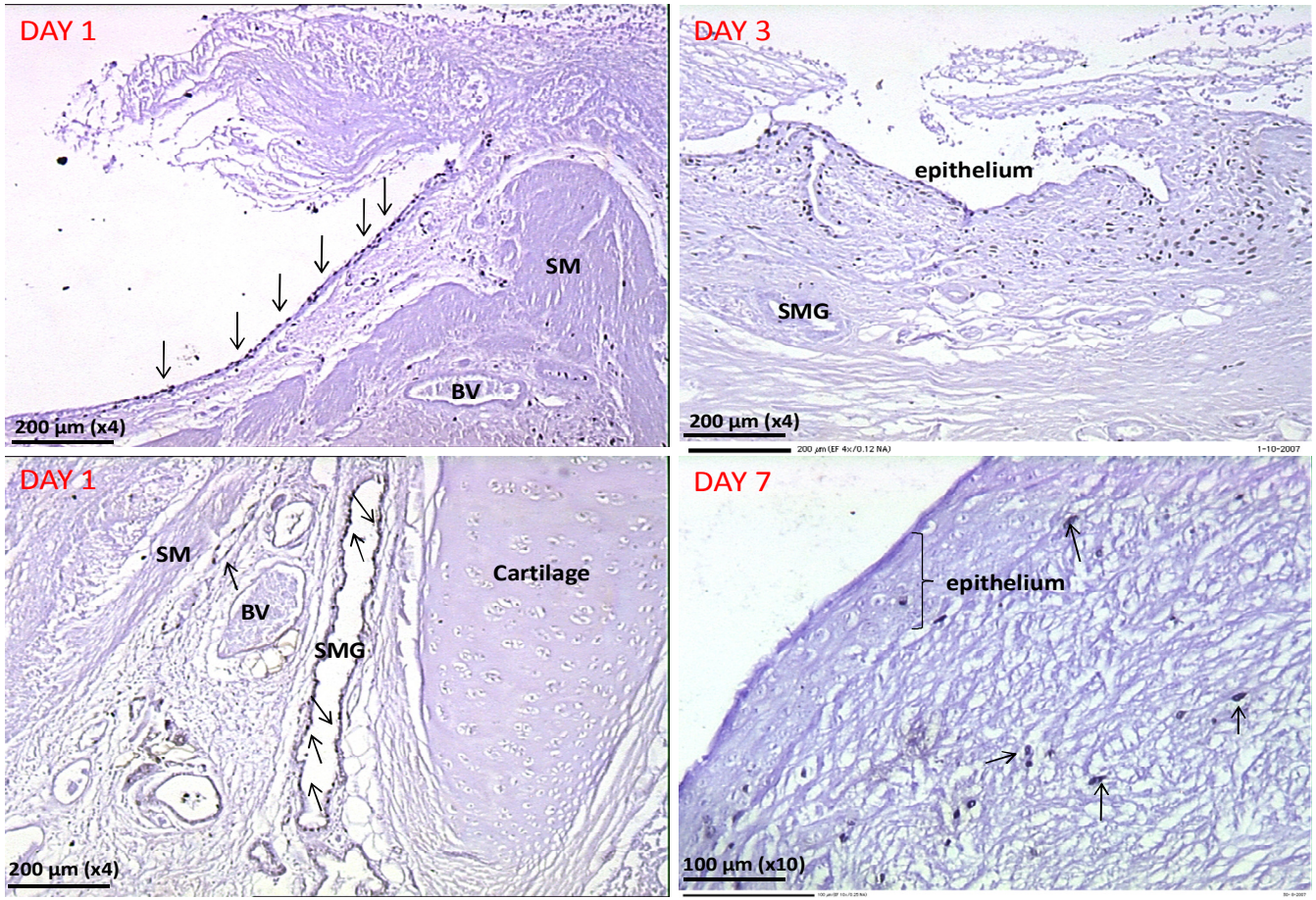


Figure 2.16: Airway tissue sections from samples harvested 1, 3 and 7 days post injury stained with BrdU

At day 1 post injury, BrdU-labelled cells were clearly visible at the transitional areas bordering either side of the lesion (top left panel). These positive cells also accumulated underneath the damaged epithelium in the mucosa and submucosa– particularly in the submucosal glands (bottom left panel). At day 3, a similar pattern of BrdU-labelled cells was observed in the epithelium and SMGs with day 1 of injury but additionally large numbers of labelled cells were present in the underlying mucosa (top right panel). At day 7, the number of BrdU-positive cells reduced particularly in the SMGs. Very few BrdU-labelled cells were present in the squamous epithelium formed over the lesion (bottom right panel). Arrows show the BrdU-labelled cells. (SMG=submucosal gland; SM=smooth muscle; BV=blood vessels).

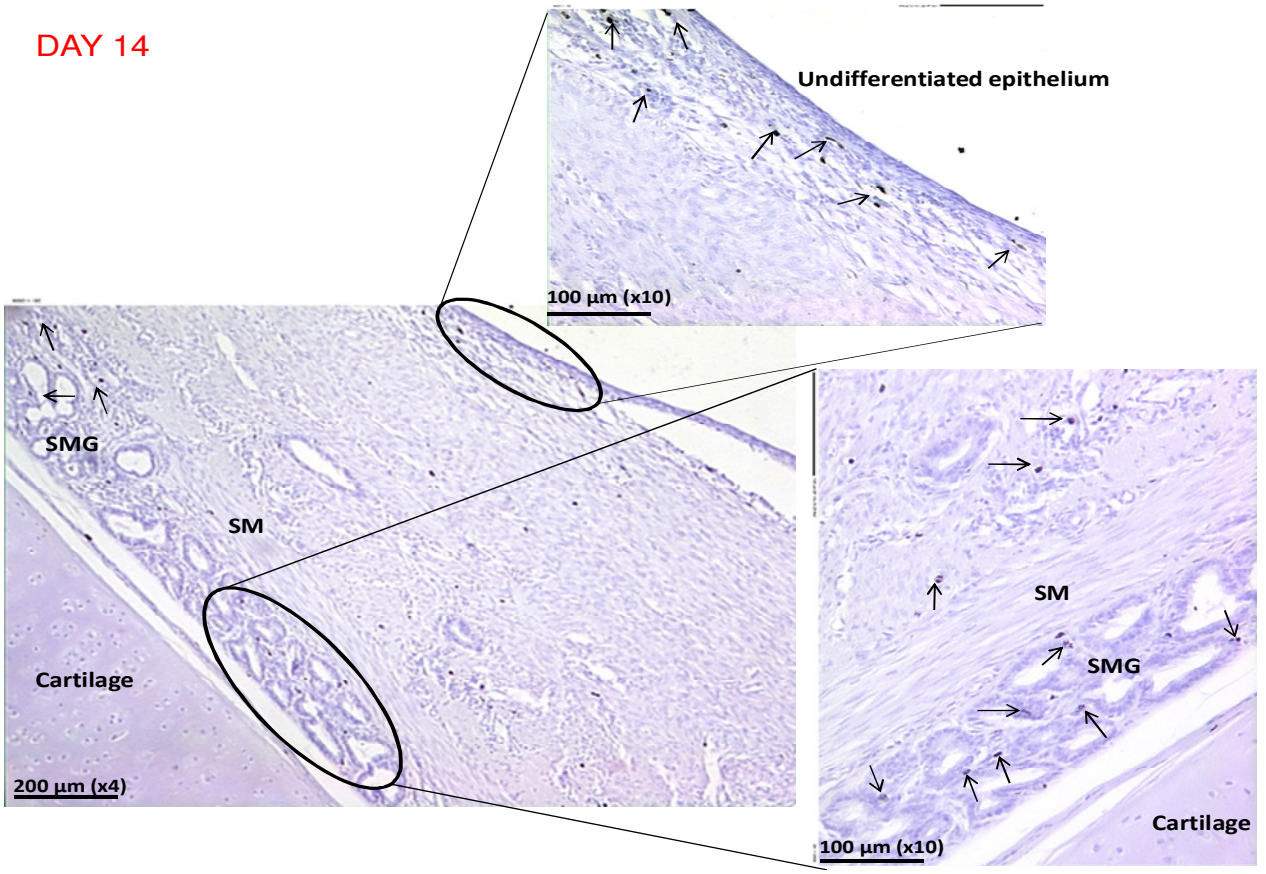


Figure 2.17: Airway tissue sections from airway samples derived 14 days post injury stained with BrdU
At this time point, although the epithelium at the damaged area consisted of partially differentiated epithelial cells, the number of BrdU positive cells was decreased relative to day 7 post injury in both the airway epithelium and in the SMGs. Arrows show the BrdU-labelled cells. (SMG=submucosal gland; SM=smooth muscle).

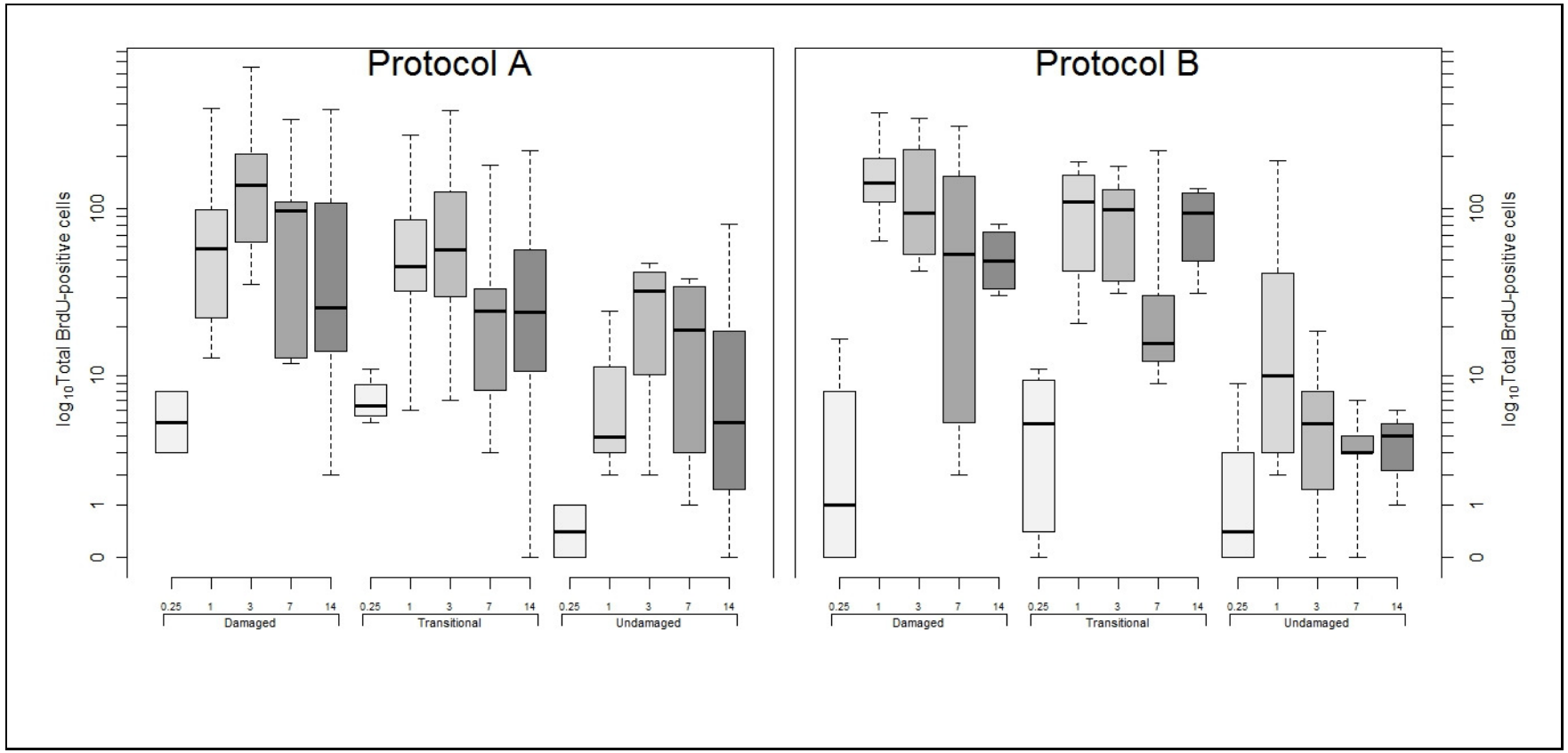


Figure 2.18: Dynamics of proliferative activity following injury
Boxplot illustrates the time-course of proliferative activity in terms of number of BrdU positive cells separated by specific areas of interest, representing damaged, transitional and undamaged epithelium, for each protocol (A and B). Upper and lower box plot margins represent the interquartile range; middle bar indicates the median. The graph was plotted based on the absolute number of cells stained with BrdU on each area measured. The points outside the ends of the whiskers are outliers.

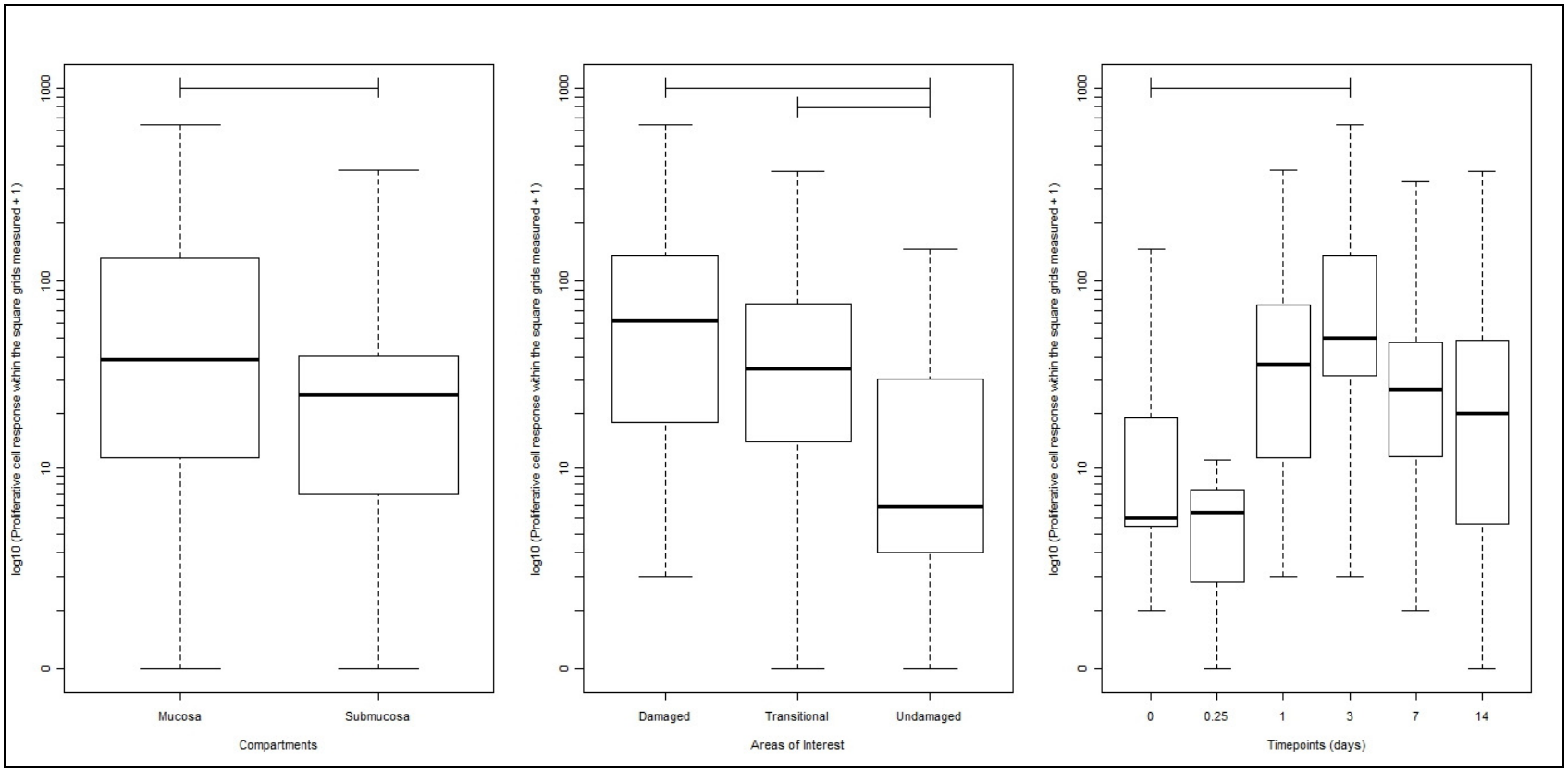


Figure 2.19: Distribution of BrdU-positive cells

Boxplot illustrating the numbers of BrdU cells counted in sections from sheep in Protocol A, separated by compartment, area of interest and time points.
Overall cell count was significantly greater for the mucosal versus submucosal compartment ($p<0.001$) and for the damaged and transitional areas of interest relative to the undamaged areas response ($p<0.001$) in both compartments. There was a significant increase in the number of BrdU-positive cells at day 3 post injury as compared to naïve (timepoint 0) ($p=0.008$). Upper and lower box plot margins represent the interquartile range; middle bar indicates the median. The points outside the ends of the whiskers are outliers. The graph was plotted based on the absolute number of cells stained with BrdU on each area measured.

2.4.2.2 CD45 expression

To study the involvement of blood-derived cells in airway regeneration and repair, the slides were stained with anti-CD45 antibody. Figure 2.20 to Figure 2.23 show the airway sections stained with anti-CD45 antibody for the undamaged epithelium, and for days 1, 3 and 7 post injury. In the undamaged airway, the CD45-stained cells were seen residing in the blood vessels and underneath the epithelium (Figure 2.20) and a few positive cells were seen in the vicinity of the submucosal glands. By 1 day post injury, there was a considerable increase in the number of CD45-positive cells observed in the damaged areas. Cells were observed at the area of epithelial haemorrhage and in the mucosa underneath the damaged epithelium (Figure 2.21) as well as around the smooth muscle of the submucosal compartment. Although there was evidence of positive cells within the epithelium bordering the lesion they were fewer in number than those observed underneath the epithelial basement membrane. On days 1 and 3, CD45 positive cells were observed both in the airway wall subjected to brushing and in the undamaged wall opposite – however numbers in the former greatly exceeded those in the latter. By day 3 (Figure 2.22) post injury CD45-positive cells were found within the repairing epithelium and in the vicinity of the submucosal glands. The number of positive-stained cells within these areas decreased from this point onwards and by day 7 the number and distribution of CD45 positive cells was similar for damaged and undamaged sites within the airways (Figure 2.23). Quantification of these CD45-positive cells was not attempted. A qualitative impression of the relationship between populations of BrdU- and anti-CD45 stained cells could be obtained through observing serial- or near-serial sections obtained at day 1 (Figure 2.24). Whilst some evidence would suggest that similar patterns of staining existed within the epithelium it was not possible to be definitive in this regard.

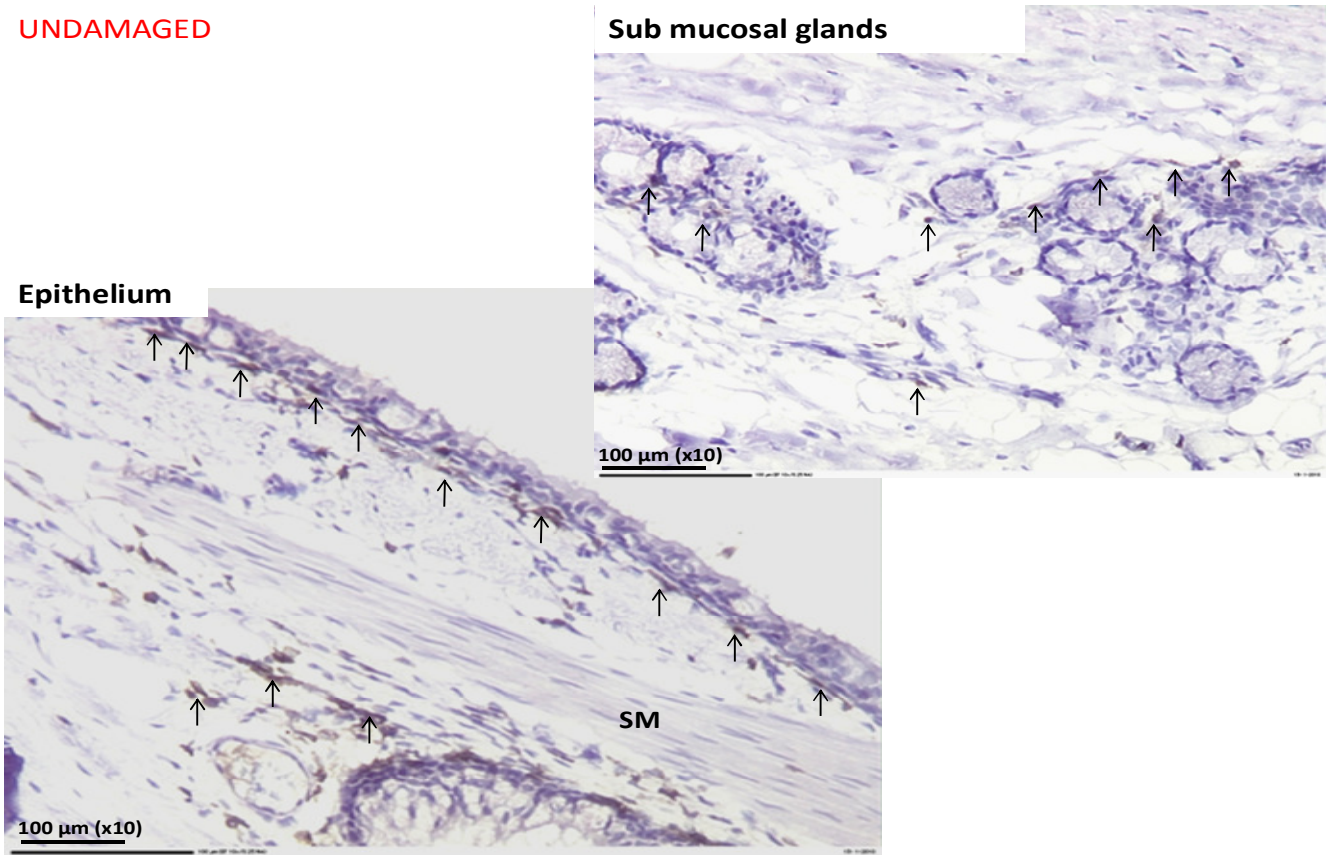


Figure 2.20: Normal airway tissue sections stained with anti-CD45 antibody
In the undamaged airway, the CD45-stained cells were seen residing in the blood vessels, underneath the epithelium and in association with the submucosal glands. (SM=smooth muscle)

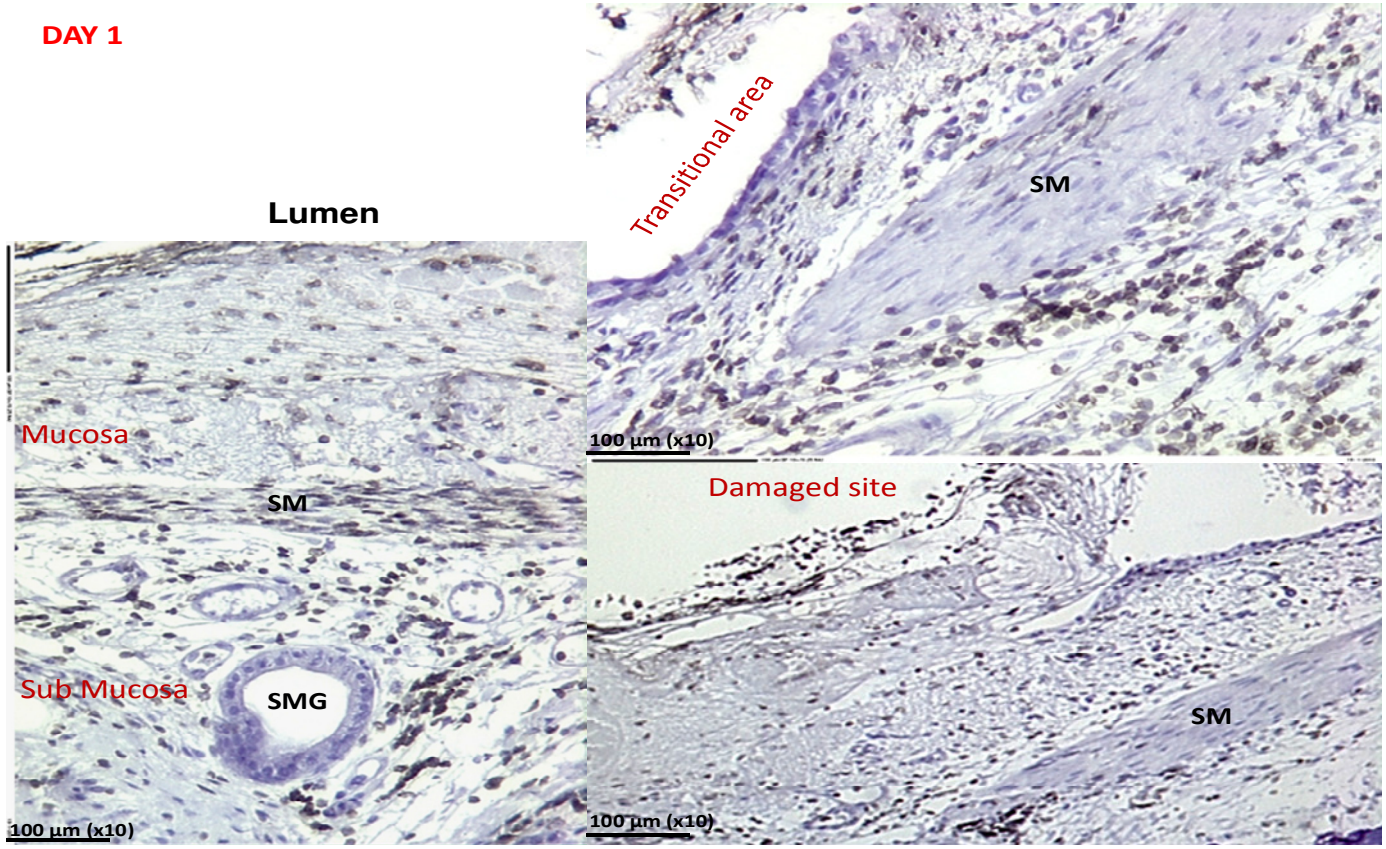


Figure 2.21: Airway tissue sections reflecting changes consistently observed at day 1 post injury (stained with anti-CD45 antibody)
At 1 day post injury increased numbers of CD45-postive cells were observed, especially at the area of epithelial haemorrhage and in the mucosal compartment underneath the damaged epithelium as well as around the smooth muscle of the submucosal compartment. (SMG=submucosal gland; SM=smooth muscle)

DAY 3

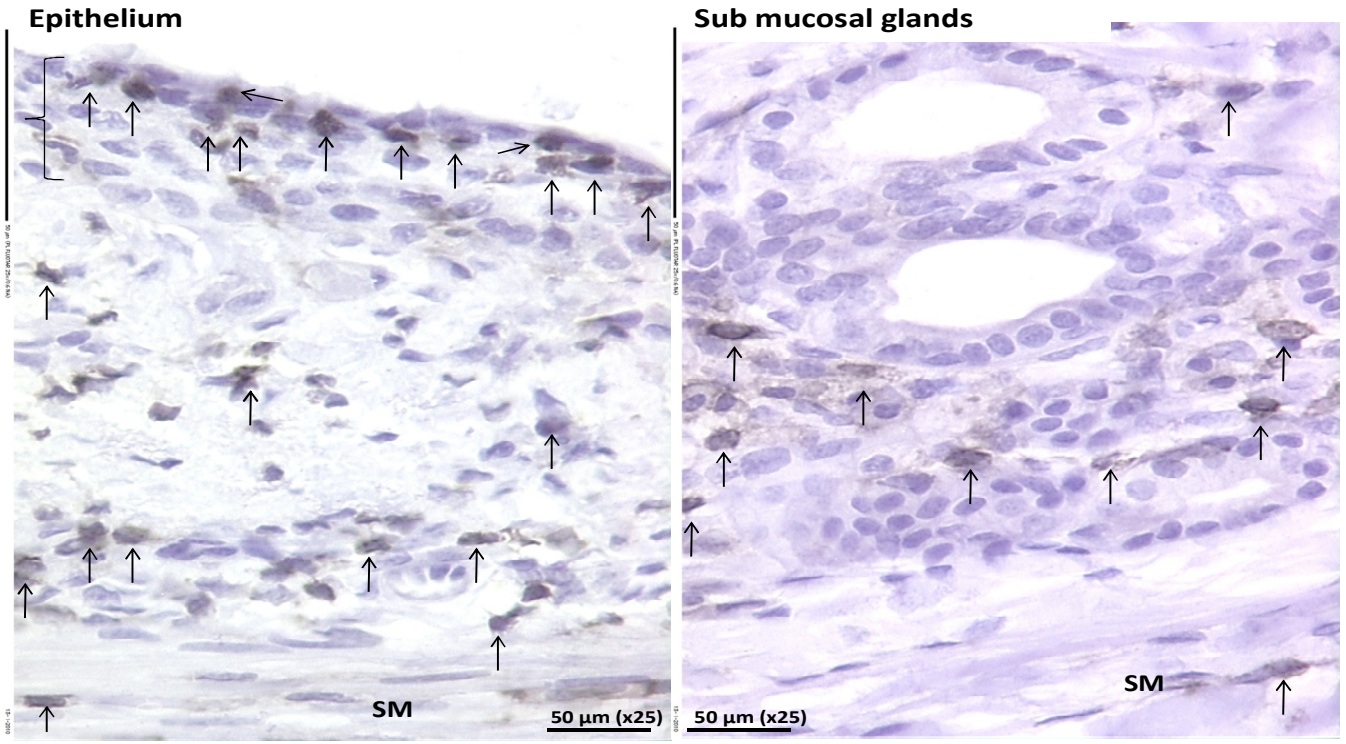


Figure 2.22: Location of CD45-positive cells at day 3 post-injury
Photomicrocraph showing CD45-positive cells (arrows) at day 3 post injury either in the repairing epithelium (left panel) or surrounding the submucosal glands (right panel)

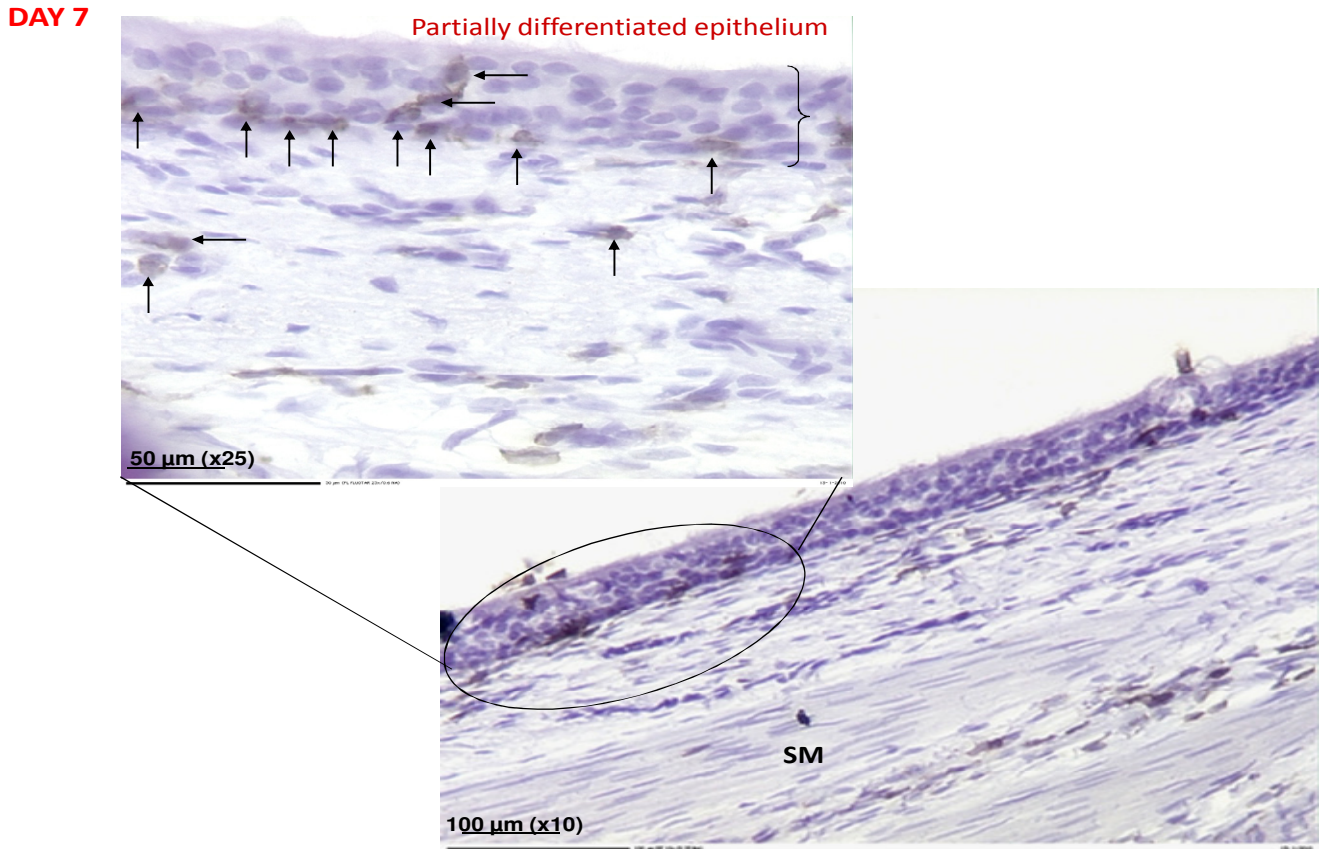


Figure 2.23: Airway tissue sections reflecting changes consistently observed at day 7 post injury
Most of the airway sections at these time points showed similar patterns of CD45 staining with the undamaged site. (SM=smooth muscle)

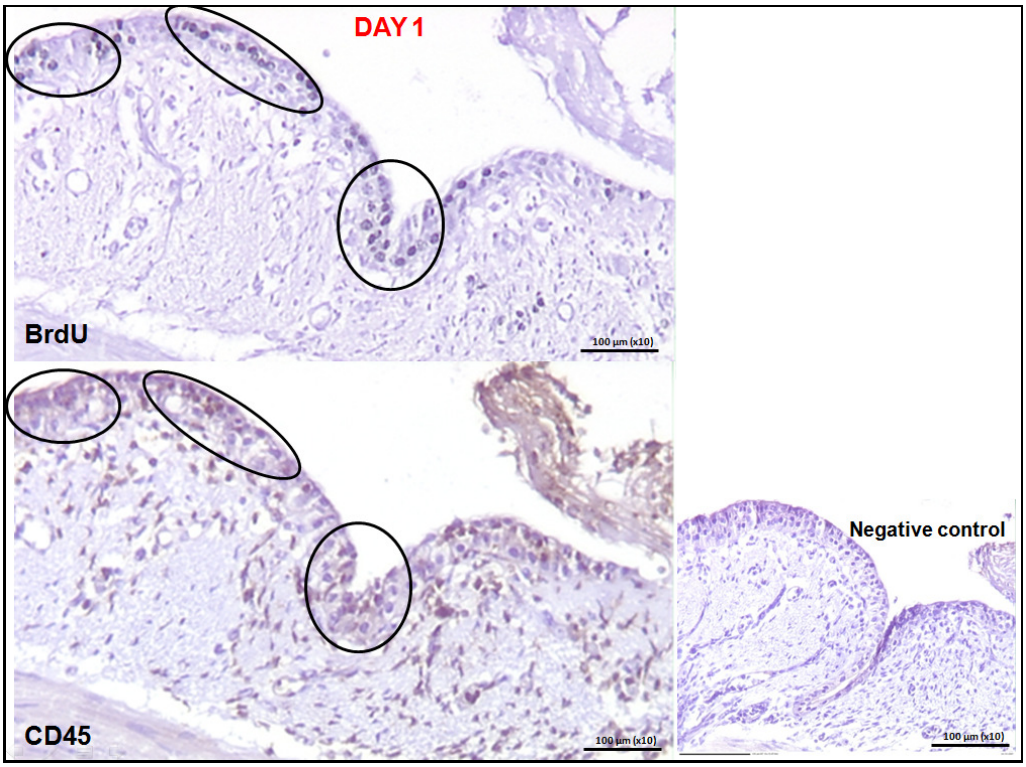


Figure 2.24: Serial sections of the airways stained with BrdU and CD45 antibodies
The circled areas suggest some evidence of co-localized CD45 and BrdU positivity within the epithelium at the margins of the injury site. A negative control section shows negative for CD45 staining.

2.4.2.3 AB-PAS staining

Slides were stained with AB-PAS as described in methods. Figure 2.25 shows images of sections from undamaged sites and brushed sites at days 1, 3 and 7 post injury. Sections derived from normal (undamaged) airways demonstrated numerous goblet cells interspersed amongst the ciliated cells that comprise the bulk of the respiratory epithelium at these locations (Figure 2.25 ; top left panel). This morphology changed dramatically following brushing, as expected where epithelial denudation occurred in the damaged area, but also within the transitional area at the wound margins. In this latter area of interest, at 1 day after brushing, the number of PAS-positive cells had markedly decreased, presumably associated with the ongoing process of epithelial flattening and dedifferentiation (Figure 2.25; top right panel).

By day 3, a layer of PAS-positive staining was observed overlying the damaged epithelial surface. Closer examination determined that the layer was comprised of a dense and compact accumulation of neutrophils held within and below the mucin layer (Figure 2.25; bottom left panel). Although an epithelial layer had reformed over the lesion at day 7 (Figure 2.25; bottom right panel) post injury, there were still very few goblet cells present within the regenerating epithelium. Those PAS-positive cells that were present demonstrated apical staining. The morphology of the regenerating goblet cells on the surface of the repairing epithelium appeared distinct from those observed in the undamaged region. At day 14 post injury, although a complete epithelial layer exists, the number of goblet cells remains lower than that observed in a normal, fully differentiated epithelium (data not shown).

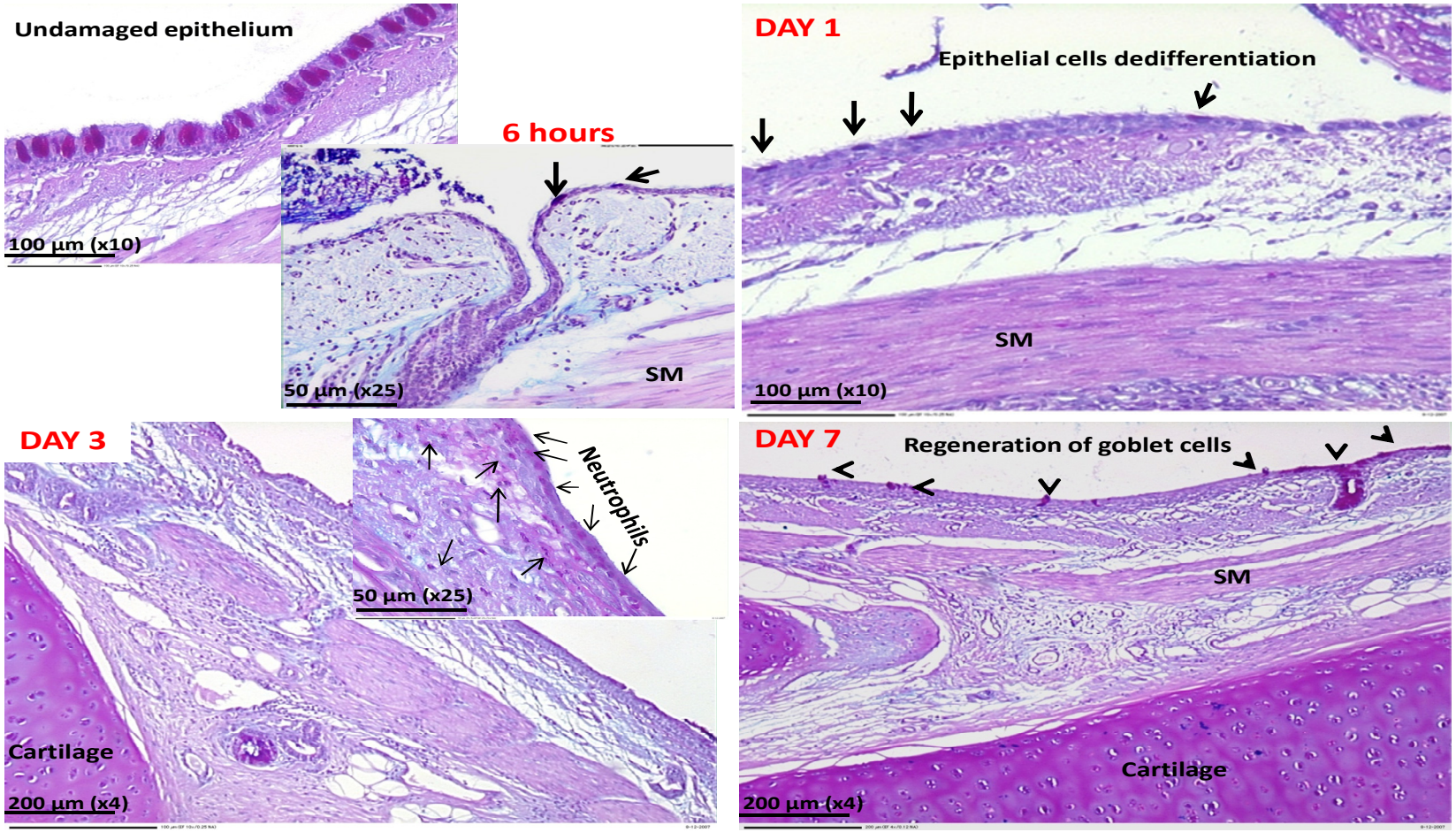


Figure 2.25: Airway tissue sections stained with AB-PAS

Figure 2.25
In the normal airways (undamaged), goblet cells are clearly stained with AB-PAS staining (top left panel). At six to twenty-four hours post-injury, PAS-positive cell is no longer retained the characteristic morphology of goblet cells but rather were attenuated in appearance (top right panel; thick arrows) which characterised as epithelial dedifferentiation. By day 3 post-injury, a layer of positive AB-PAS staining was observed overlying the repairing epithelial surface. Closer examination determined that the layer was comprised of a dense and compact accumulation of neutrophils held within and below the mucin layer (bottom left panel; thin arrows). By day 7 post-injury regenerating goblet cells could be seen incorporated in the epithelium overlying the brushed site (bottom right panel; arrows head) which characterised as epithelial redifferentiation. (SM=smooth muscle)

2.5 DISCUSSION

We chose the sheep as our model system to examine the airway epithelial response to physical injury for several reasons. Sheep are a well accepted model for lung injury and disease (Bischof RJ et al., 2003, Scheerlinck JP et al., 2008). The airway epithelial structure of the tracheobronchial tree has been well characterised and indicates good homology with the human lung (Becci PJ et al., 1978, Kennedy AR et al., 1978), the size of the sheep lung facilitates the application of bronchoscopic techniques that are equivalent to those used in human clinical respiratory medicine (Romagnoli M et al., 1999) and finally segmental approaches can be adopted wherein the response of different lung regions can be compared and contrasted over time within the same animal (Collie DD et al., 2001, Emerson M et al., 2003, Collie DD et al., 2006). This offers a significant advantage with respect to achieving statistical power in experimental design, and reducing the number of animals required to test given hypotheses.

Our experimental protocol capitalises on these latter advantages and in addition allows for the phased induction of lung injury in order to assess the time-dependent changes that occur during the repair process. To our knowledge this is the first time such a system has been employed in the context of understanding the response to physical lung injury in any mammalian system. In addition, the use of BrdU labelling, whilst previously used in sheep to assess cell proliferation in wool follicles (Holle and Birtles, 1990) has not been previously used to follow similar phenomena in the lung.

Accepting the novelty of the model system and protocol design, constructive critique could be directed towards individual aspects. One aspect relates to whether the selected injury sites were chosen from roughly equivalent levels in the tracheobronchial tree. It is known that proximal and distal airways do not share the same cell composition (Plopper CG and Hyde DM, 2008) and are differentially affected by pathology in diseases such as asthma and COPD (Jeffery PK, 2001) therefore it was considered important to assess whether the major aspects of structure were equivalent between airways selected from different lung segments. In this study

the results confirmed that the relative proportions of different tissue components did not differ significantly between segments. A second aspect relates to the reproducibility of the bronchial brushing technique. Whilst every effort was made to ensure consistency between sites, such that the same numbers of brush passes were employed and equivalent areas exposed to injury, it is impossible to definitively ascertain that the injuries were exactly the same. This contrasts to some extent with previously reported injury systems in surgically prepared small animals (Kim JS et al., 1997a, Wilhelm DL, 1953, Keenan KP et al., 1982a, Keenan KP et al., 1982b, Keenan KP et al., 1982c, McDowell EM et al., 1979, Shimizu T et al., 1994), or even to *in vitro* (Zahm JM et al., 1991) systems where the nature of the preparation allows for greater precision in the nature and extent of the injury process.

Initial protocols were designed to follow the response over 35 days post injury, reflecting uncertainty over the likely duration of the repair process in this species. Later time points were subsequently dropped since initial studies indicated that in the majority of cases recovery had occurred by 21 days post injury. Since our principal interest was in understanding the mechanisms underlying the normal processes of repair in the airway and potentially relating those findings to situations where the airway is continually faced with insult we felt that the early processes, involving significant cell proliferation, would likely yield the most useful information to compare against a chronic setting.

Of interest is the observation that the brushed areas were covered by partly differentiated epithelial cells as early as day 7 post injury. In the study of (Heguy A et al., 2007) involving healthy human subjects, the damaged airways regained their normal structure by day 14 post injury – by day 7, the damaged areas were similarly covered by undifferentiated epithelial cells. Clearly whilst the model systems are not exactly equivalent, it is of some interest that the repair response to similar injury appears to follow a similar temporal spread between these two species.

It was clearly demonstrated that the different segments of the lung share similar airway composition and structure. This reduces the possibility that any differences

between sites might reflect differences in airway structure, related to differences in the position in the airway tree. Although this was only a preliminary study with low numbers, it appeared that the available data supported the notion that the pattern of changes in response to physical injury was similar whether the changes were assessed over time within the same animal (protocol A) or were assessed between different animals, with only one time point represented per animal (protocol B). Although it is often assumed, it should not be concluded that spatial separation allows for independence of treatments. Indeed, in previous studies by (Heguy A et al., 2003), examining the early transcriptional response to bronchial brushing in the lung, it was demonstrated that early changes can be observed in the right lung if sampled after a preceding procedure in the left lung and vice versa, depending on the sequence and timing of individual samples.

Protocol B was limited in terms of there being only one sheep per time point and the possibility remains that an individual sheep effect might have contributed to the observed difference in the timing of the peak proliferative response subsequent to injury between the two protocols. For this reason interpretation is necessarily limited to a qualitative appraisal of the general patterns of response, rather than influenced by specific comparisons at certain time points.

In the normal state, stem cells have a slow cell turnover, giving rise to transient amplifying (TA) progenitor cells that retain the capacity to replicate and impart most of the tissue renewal in the presence of injury. However, in the absence of injury or cell renewal, TA cells eventually become incapable of proliferation and become terminally differentiated (TD) cells (Borthwick DW et al., 2001). The slow cycling feature of stem cells lends itself to the use of DNA labelling techniques followed by a “washout” period to track candidate stem cells *in situ* (Liu X and Engelhardt JF, 2008, Borthwick DW et al., 2001) as “label-retaining cells”. In this study a single dose of BrdU, which incorporates during S-phase of the cell cycle (Yanai T et al., 1996),with no washout period was used to investigate the presence of cells with proliferative activity. The BrdU-labelled cells observed in the small intestine and skin sections indicate that the dose of BrdU used in this study was appropriate, as

well as validating the efficacy of the BrdU infusion protocol. The choice of these tissues as the source of control material was based on the study carried out by (Holle and Birtles, 1990) for the skin, and by (Yanai T et al., 1996) for the intestine. The patterns of staining observed in these tissues was consistent with that anticipated from these publications.

Although a cellular response to physical injury was seen as early as 6 hours post injury, this response was apparently not accompanied by a significant increase in proliferative activity. The observed cellular response at this time point was believed to reflect preliminary morphological changes involving epithelial cells, and would not necessarily involve significant proliferative activity. Histopathological evidence suggests that the observed epithelial attenuation and cellular migration was devoted towards the initial regeneration of a single layer of squamous epithelial cells that would cover the denuded surface. It is conceivable that the purpose of this migratory process is to provide a foundation for other cell types migrating from either local epithelial cell population (Hong KU et al., 2004b, Hong KU et al., 2004a, Rawlins EL et al., 2007, Reynolds SD et al., 2000) or from the circulation (Abe S et al., 2004, Anderson DJ et al., 2001, Kotton DN et al., 2001) in order to contribute to the repair process. This assumption is based on the discussions by Lane BP and Gordon R, (1974), and (Rennard SI et al., 1991) which describe early peripheral cell migration in order to cover the tracheal epithelial defect induced by mechanical trauma. Although there was evidence of neutrophil infiltration at the damaged area, these were fewer in number when compared to day 1 post injury.

Whilst based only on scrutiny of one segment, preliminary data indicates that secretory or mucus-producing cells are involved in the repair process from early as 6 hours post injury. The involvement of PAS-stained cells in the early repair process was reported by (Shimizu T et al., 1994) who described these cells migrating from the undamaged epithelium bordering the lesion following a mechanical injury in rats. These cells were characterized as partially differentiated cells due to the loss of the normal differentiation markers of mature epithelial cells.

The histopathological features of the airways at day 1 differed to those observed at day 3 post injury where most of the damaged areas were clearly visible with mucosal haemorrhage and most of the BrdU-positive cells were clearly seen at both transitional areas and underneath the damaged epithelium. At this time point, the epithelium at the transitional area was thickened with multiple layers of undifferentiated cells. By day 3 post injury the brushed area was covered mostly by intact epithelium with some areas thickened as a consequence of hyperplasia. This finding appeared to parallel the observed peak in cell proliferation at the same stage. This result somewhat contrasts with findings in a guinea pig tracheal injury model in which proliferative activity peaked at 48 hours and began to decrease at 72 hours post injury (Kim JS et al., 1997a).

The timing of the peak proliferative response might be associated with the severity of the brushing technique and respective size of the lesion. It is likely that the technique performed in this study would be classified as severe (fourth category, with no intact basal cells left following injury (Barrow RE et al., 1992)) since the cellular composition of the airway epithelium was removed completely, and the basement membrane and various elements of the mucosal structures were also disturbed or destroyed. This contrasts with many of the small animal systems where a layer of viable basal cells is left at the wound site, apparently facilitating the migration process and promoting early proliferation and repair (Keenan KP et al., 1982a). The impact of more severe damage on the delay of the proliferation peak was also reported by others (Keenan KP et al., 1982a, Keenan KP et al., 1982b, Keenan KP et al., 1982c, Lane BP and Gordon R, 1974, McDowell EM et al., 1979, Wilhelm DL, 1953). Severe injury is thought likely to disturb cell-to-cell communications during the early period of the injury (Keenan KP et al., 1982a) and disrupt the basement membranes in the denuded areas (Keenan KP et al., 1982c) - both believed essential to the early process of epithelial cell migration. Indeed the basement membrane provides a mechanical support for cells, acts as a semi-permeable barrier between tissue compartments and regulates cellular attachment, migration and differentiation (Schittny and Yurchenco, 1989).

Tissue sections from day 1 to day 3 post injury demonstrated that the newly regenerated epithelial surface contained some positive AB-PAS staining. The AB-PAS stain also nicely highlighted the dedifferentiation process of cells in the transition areas adjacent to the injury sites. Whether mucus-producing and/or ciliated cells contribute most to the processes of dedifferentiation and migration is unknown in the context of the current study. However, in a study of mice exposed to naphthalene it was suggested that re-differentiation of the bronchiolar epithelium after naphthalene injury involves spreading and trans-differentiation of ciliated epithelial cells with concomitant changes in cell shape and gene expression during the process of repair (Park KS et al., 2006). Thus, our finding might support the hypothesis that either ciliated and/or mucus-producing cells play a role in contributing to the repair of the epithelium exposed to physical injury either through dedifferentiation (Erjefält JS et al., 1995) or trans-differentiation (Park KS et al., 2006) and migration towards the damaged area. The involvement of the cells bordering the lesion was concurrent with the peak in proliferative activity. The high numbers of BrdU-positive cells observed at the transitional area compared with those in the undamaged area indicate that the repair process involves neighbouring cell populations not directly overlying the area exposed to the injury. This phenomenon was clearly described by (Keenan KP et al., 1982a, Keenan KP et al., 1982b, Keenan KP et al., 1982c) in which the early repair events involve cell spreading and migration into the denuded area.

It is interesting to note that there was also a significant involvement of the submucosal glands underlying the wound, with evidence of proliferation in the gland epithelium. The increased number of BrdU-labelled cells in the epithelium and submucosal glands of the airways might suggest that differentiation of mucus secreting progenitor cells and their subsequent proliferation might be a primary aim of the healing process. Indeed, an intact mucus ‘scaffold’ overlying the damaged area observed at day 1 to day 3 post injury might be a pre-requisite for successful healing – preventing further damage, whilst facilitating the migration, attachment and interaction of the cells participating in the healing process. The reason for this response in the glands is unclear, but could reflect an increase in gland secretory

activity to help in forming a protective barrier over the wound site. There is a substantial amount of literature on SMG secretions as a source of significant levels of antimicrobial molecules and host defence i.e lactoferrin (Lf), lysozyme (Ly). In a study by Raphael GD et al., (1989) to investigate the cellular sources and the secretory control of these nasal proteins in vivo adults underwent nasal provocation tests with methacholine (MC), histamine (H), and gustatory stimuli, reported that Lf and Ly, two antimicrobial proteins previously described in nasal secretions, are produced and stored in the nasal submucosal glands, specifically within serous cells, and are secreted under the influence of the parasympathetic nervous system (Raphael GD et al., 1989). This study demonstrated that these molecules are important constituents of nasal secretions (Raphael GD et al., 1989). The secretion of these molecules by the submucosal glands support the concept that secreted mucus serves as a lubricant (Raphael GD et al., 1989) for the airway epithelium surface as well as in agreement with predominant role of submucosal gland in the upper airways to provide a first line of defence for the mucosa (Raphael GD et al., 1989). In addition, it was also been suggested that lactoperoxidase (LPO) is an important component of airway defence against infection and inhibiting the LPO system in sheep impaired bacterial clearance from the airway (Gerson C et al., 2000). Thus it was suggested that LPO system is a major contributor to airway defences (Gerson C et al., 2000).

Mucus production following airway injury might be relevant to the potential protective role, or it could also be the response of the gland to initiate the repair of duct structures damaged by the brushing. This hypothesis is supported by the observation that production and secretion of the of airway mucus from the mucous goblet cells and other secretory cells within the surface and underlying submucosal glands are significantly increased in human diseases such as cystic fibrosis, chronic bronchitis and asthma (Wang K et al., 2007, Engelhardt JF et al., 1995). It is also tempting to speculate that the proliferative activity seen in the submucosal glands underneath the damaged mucosa might in some way reflect the fact that a stem cell niche exists in this vicinity (Engelhardt JF, 2001) and such proliferation might reflect the need to replace airway epithelial cells.

The concordance, in terms of location, between the BrdU staining cells at the margins of the lesion, and in the submucosal gland structures, is intriguing and suggests that a major focus of the cell proliferation could be directed towards early (day 1) migration and regeneration of the epithelium at the lesion. Increased AB-PAS staining and BrdU-labelled cells at the submucosal gland region at this time point suggest early production of mucus to act as a protective barrier to avoid further damage to the epithelium. This assumption is supported by the evident of the layer comprised of a dense and compact accumulation of neutrophils held within and below the mucin layer. Whilst the specific lineages of the proliferating airway epithelial cells and where they originated or migrated from are still unclear the literature suggests several possible sources including migration from the areas bordering the damaged site, from submucosal glands, or perhaps even involvement haematopoietic-derived cells. In the latter regard we sought to assess the contribution of blood-derived cells in the repair process through anti-CD45 staining. Given that CD45 positive cells can be observed within the airway epithelium bordering the lesion, one could hypothesise that these cells might also be proliferating since there was an abundance of BrdU-positive cells observed in this location, especially at day 1 post injury. However, only a few CD45-positive cells were observed within the repairing epithelium or within or around the glands in the submucosal region. These cells were scattered in both areas of mucosal and submucosal compartment and also observed in regenerating smooth muscle (day 1). However, it seems unlikely that CD45 positive cells account for all proliferative activity in these areas since the numbers of cells stained with anti-BrdU in the transitional epithelium exceed the numbers of CD45 stained cells. A dual-labelling technique might be a promising tool in order to demonstrate this hypothesis; however the protocols for these antibodies are not compatible – both antibodies share a similar IgG group (IgG_1). Of related interest is the observation that in an experimental sheep model of allergic lung inflammation after local challenge with house dust mite, it was revealed that CD45-positive cells were the most prominent leucocytes found in lung tissue 48 hours after allergen challenge (Bischof RJ et al., 2003). Naturally, the perceived roles of such cells can extend to a range of influence including the secretion of chemokines and cytokines essential for the repair process (Chuang CY et al., 2009).

The AB-PAS staining at day 3 post injury was arranged in a thin covering over the top of the injured epithelial surface, and this layer was packed with neutrophils (Figure 2.25). Whilst it is well recognised that neutrophil products promote mucin production (Kim S and Nadel JA, 2004) the observed relationship between such induction and the cells held within the layer perhaps hints at some functional benefit of having these infiltrating leukocytes constrained in such tight apposition to the repairing epithelium. Indeed neutrophils have important phagocytic functions that might relate to the need to remove dead or dying cells and cell debris within the surface and equally might serve an important additional protective function in bolstering innate defence at this site. In such light it is also tempting to speculate that the observed neutrophilic influx that often characterises exacerbations in chronic airway diseases such as COPD has its origins in repair rather than inflammation *per se*.

In the present study, the number of BrdU-positive cells decreased by day 7 post injury. By this time, although the epithelial surface was fully covered, some areas still consisted only of partially differentiated epithelial cells. The reduction in BrdU-incorporating cells at day 7 may indicate that at this time point, the majority of the proliferating cells might have entered the late stage of cell cycle, and fail to label with BrdU which is only incorporated during the S-phase of the cell cycle. In the study by (Heguy A et al., (2007) most of the genes picked up on microarray analysis at day 7 post injury were the genes involved in the late stage of the cell cycle.

Re-differentiation of the epithelium in the form of increased AB-PAS-positive staining was clearly evident by day 7 after injury. Whilst these cells were clearly differentiating towards a mucus-secreting phenotype they were still morphologically distinct from the fully differentiated mature goblet cells present in the undamaged epithelium. Some concordance was evident in the location and extent of PAS-staining and proliferative activity present in the submucosal glands. Whether such proliferation contributes to increased airway secretions and/or functions as a stem cell niche to promote airway epithelial renewal, remains to be determined (Liu X and Engelhardt JF, 2008).

Basal cells are thought to make a major contribution to the processes of migration, proliferation and differentiation during regeneration of the damaged proximal airway epithelium (Hong KU et al., 2004b, Hong KU et al., 2004a). These cells may have originated from the area bordering the injured epithelium, by changing their shape (Keenan KP et al., 1982a, Keenan KP et al., 1982b, Keenan KP et al., 1982c, Kim JS et al., 1997a, Wilhelm DL, 1953, Zahm JM et al., 1991) through a dedifferentiation process and becoming intermediate cells which then give rise to other specialized epithelial cells (i.e ciliated cells). Contrary evidence suggests that secretory cells are the principal progenitor cells in the regenerating tracheal epithelium of hamsters and rats (Keenan KP et al., 1982a, Keenan KP et al., 1982b, Keenan KP et al., 1982c, Lane BP and Gordon R, 1974, McDowell EM et al., 1979). These latter studies also demonstrated that such secretory cells provide the majority of the cells that change shape and migrate to cover the wound and divide to produce a multi-layered metaplastic epidermoid epithelium. Progeny derived from such cells developed into columnar secretory cells (pre-secretory cells) or columnar ciliated cells (pre-ciliated cells) in order to generate a pseudostratified mucociliary epithelium (Keenan KP et al., 1982a, Keenan KP et al., 1982b, Keenan KP et al., 1982c). The scope of the present study was such that it was not possible to ascertain the respective roles of either cell type in airway regeneration and repair in the sheep.

Although the preceding discussion has largely focused on the changes observed in the damaged and transitional areas it was notable that changes were also observed in the undamaged area opposite to the brushed area. The result was similar to the finding reported by (Kim JS et al., 1997a) in the guinea pig tracheal epithelium (Kim JS et al., 1997a) in which a proliferative response was observed in the undamaged area opposite the damaged area away from the damaged site.

2.6 CONCLUSION OF THE CHAPTER

In conclusion, this study demonstrates the utility of a novel model of airway epithelial injury in a large animal model system that can be used to follow the repair process. The model demonstrated changes in airway architecture and histopathology largely consistent with previous studies in small animal systems and with similar studies conducted in man. The processes of dedifferentiation, migration, proliferation and re-differentiation were easily characterised and assessed by selectively injuring areas of the airway epithelium. The model further substantiates the notion that the airway epithelium demonstrates remarkable propensity to repair itself in the event of acute injury. Indeed, although the brushing technique used in this study destroyed most of the morphology of the mucosal region, the ability of the airway to regenerate a new partly differentiated epithelium and other airway components such as smooth muscle and glands as early as 3 to 7 days post injury was remarkable. The chapter also demonstrates the potential of experimental protocols involving segmental challenges to maximise data yield whilst minimising the ethical costs associated with particular questions relating to airway epithelial repair.

CHAPTER 3

A CANDIDATE GENE APPROACH TO ANALYSING GENE EXPRESSION DURING AIRWAY REGENERATION AND REPAIR FOLLOWING INJURY

3.1 INTRODUCTION

Understanding the mechanisms involved in the response to injury in the lung is fundamental to the process of piecing together the pathways and mechanisms that underlie more complex respiratory diseases such as asthma and COPD. Amongst the various mechanisms, pathways regulating lung morphogenesis may play an important role in shaping the nature and extent of tissue remodelling that characterises the pathology associated with these conditions. We are interested in determining the default process of repair that is evoked when the airway epithelium is exposed to physical injury, and also whether this process has its origins in the developmental biology of the lung. Indeed, our working hypothesis is that many of the mediators and pathways that are known to play an important role in lung development are also involved in the repair and remodelling of the airways in response to injury. A number of studies have linked lung development to concomitant changes in gene and protein expression (Costa RH et al., 2001, Warburton D et al., 2000). Whilst the process of wound healing has been examined in the airways grossly and at light and electron microscopy resolution (Shimizu T et al., 1994, Erjefält JS et al., 1997, Erjefält JS et al., 1995) there is limited information available on the changes in gene expression associated with repair.

Lung development is comprised of a series of interrelated morphogenetic events. Kho AT et al., (2009) have identified the molecular signature of the lung during the transition towards air breathing – a period they refer to as “time to birth”. Amongst the statistically overrepresented genes were those involved in oxygen and gas transport activity, as well as host defence function and those concerned with the initiation of alveolar formation (Kho AT et al., 2009). In addition some more unexpected genes were identified including those connected with nitric oxide synthesis, and ribosomal and haemoglobin complex formation. These genes are informative in terms of defining the relative state of mouse lung tissue in terms of its preparedness for the transition to air breathing, and thus represent a molecular signature that effectively describes the distance in age from the day of birth (Kho AT et al., 2009). More recently another study by Kho AT et al., (2010) characterized global patterns of gene expression occurring during human lung development and

similarly indicated the existence of molecular phases of lung development (Kho AT et al., 2010). In early pseudoglandular development, genes were enriched with chromosomal organization processes associated with mitosis, and the late pseudoglandular genes were enriched with surfactant function–gas exchange and immunological-major histocompatibility complex (MHC) class II attributes (Kho AT et al., 2010) with surfactant protein genes increased from the early to late pseudoglandular phase. The early expression of surfactant proteins may indicate a necessity of early molecular programming of the lung for subsequent development (Kho AT et al., 2010).

These findings concur with opinion expressed in the review of (Warburton D et al., 2001) in which it was acknowledged that modelling of the lungs during morphogenesis, repair, and regeneration is tightly coordinated by conserved stimulatory and inhibitory signalling mechanisms, including specific transcriptional factors, cytokines, peptide growth factors, proteases, and matrix elements (Warburton D et al., 2001). Genes that have emerged from functional screening studies as directing fly respiratory organogenesis, including sonic hedgehog, patched, smoothened, Gli, fibroblast growth factor (FGF), fibroblast growth factor receptor (FGFR), sprouty, bone morphogenetic protein 4 (BMP4), TGF-beta and SMAD, are functionally conserved in mice and humans(Warburton D et al., 2001).

It is clearly important for the process of repair to be highly regulated, given the complexity of lung structure and the crucial interplay between structure and function in this organ. Therefore, it is likely that default pathways do indeed exist and are evoked in response to acute injury, and that it is only when such pathways are perturbed or overwhelmed that the sequences of events result in pathology. Successful repair will likely involve a series of interlinked and interdependent events. Pathology may occur where aspects of sequence and/or interdependence are compromised and knowledge of where such abnormalities arise offers the potential of modulating repair thus minimising the chance of developing significant lung pathology. Although the complexity of airway epithelial cell phenotypes and their

involvement during injury and repair is well discussed, there is a lack of evidence and detailed explanation of how the repair process is regulated.
Cell migration and proliferation have been shown to be important in wound healing in the airway (Chapter 1 Section 1.6.1) and such activities i.e cellular proliferation, migration, and differentiation, also play key roles during branching morphogenesis in lung development (Costa RH et al., 2001). However, the factors that elicit and control such processes have not been fully determined. Transcription factors, such as transcription termination factor 1 (TTF-1), fox family members (including Foxa1, Foxa2, and FoxJ1), GATA-6, and the beta-catenin/TCF transcriptional complex, influence genetic programmes critical for lung morphogenesis, differentiation, and pulmonary homeostasis (Costa RH et al., 2001). These transcription factors control the expression of genes that are critical for the differentiation of the distinct subsets of cells characteristic of the mature lung and act as mediators for epithelial-mesenchymal interactions during lung development and lung branching morphogenesis (Cardoso WV and Lü J, 2006, Costa RH et al., 2001, Mucenski ML et al., 2003, Maeda Y et al., 2007).

Beta-catenin is part of a complex of proteins that constitute adhesion junctions which are necessary for the creation and maintenance of epithelial cell layers by regulating cell growth and adhesion between cells. Beta-catenin also anchors the actin cytoskeleton and may be responsible for transmitting the contact inhibition signal that causes cells to stop dividing once the epithelial sheet is complete (Rawlins EL et al., 2007). The expression of this gene in respiratory epithelium of the embryonic lung was required for the growth and differentiation of peripheral epithelial cell progenitors (Mucenski ML et al., 2003). Beta catenin is involved in the canonical Wnt pathway and canonical Wnt ligands act extracellularly to promote the stabilization of beta-catenin in the cytoplasm (Nusse R et al., 2008). The Wnt terminologywas originally derived through a combination of Wg (for wingless) and Int (for integration) (Rijsewijk F et al., 1987). Wg was originally identified as a recessive mutation affecting wing and haltere development in Drosophila melanogaster (Sharma RP and Chopra VL, 1976) whilst Int genes, the vertebrate homologues, were initially recognised as being near several integration sites of

mouse mammary tumor virus (MMTV) (Nusse R et al., 1984). It was suggested that beta-catenin functions to control the balance between progenitor expansion and epithelial differentiation that is required for both lung development and repair after injury (Bellusci S, 2008). It was also demonstrated that Wnt signalling is activated upon lung epithelial regeneration and increased Wnt signalling expands the numbers of bronchioalveolar stem cells (BASCs) (Zhang Y et al., 2008).

In the context of lung development, De Langhe SP and Reynolds SD (2008) discussed the involvement of this gene in Wnt signalling during lung organogenesis. Wnt/beta-catenin signalling regulates aspects of branching morphogenesis, regional specialization of the epithelium and mesenchyme, and establishment of progenitor cell pools (De Langhe SP and Reynolds SD, 2008). The beta-catenin and the Wnt/beta-catenin signalling pathways contribute to the control of cellular proliferation, differentiation and migration (De Langhe SP and Reynolds SD, 2008). In contrast, Zemke AC et al., (2009) reported that beta-catenin was not necessary for maintenance or efficient repair of the bronchiolar epithelium. In a study of transgenic and cell type-specific knockout strategies to determine roles for beta-catenin-regulated gene expression in normal maintenance and repair of the bronchiolar epithelium it was demonstrated that repair of the naphthalene-injured airway proceeded with establishment of focal regions of beta-catenin-null epithelium (Zemke AC et al., 2009).

A further study demonstrated that beta-catenin, Foxa2, FoxJ1 and Sox family members (Sox17 and Sox2), are dynamically regulated during repair and re-differentiation of the bronchiolar epithelium after naphthalene injury (Park KS et al., 2006). During murine lung development, Hackett BP et al., (1995) reported that transcript for hepatocyte nuclear factor-3/forkhead homologue 4 (HFH-4/FoxJ1) was first detected at the start of the late pseudoglandular stage (embryonic day 14.5) and then was specifically localized to the proximal pulmonary epithelium (Hackett BP et al., 1995). The unique temporal and spatial pattern of HFH-4 gene expression in the developing lung defines this protein as a marker for the initiation of bronchial epithelial cell differentiation and suggests that it may play an important role in cell

fate determination during lung development (Hackett BP et al., 1995). Blatt EN et al., (1999) demonstrated that HFH-4/FoxJ1 expression is localized to ciliated epithelial cells and was temporally related to ciliogenesis. A role for FoxJ1 in ciliogenesis was demonstrated by the temporal relationship of HFH-4 to the appearance of cilia in developing lung and flagella during spermiogenesis (Blatt EN et al., 1999). It was also demonstrated that among epithelial cells expressing FoxJ1 mRNA, FoxJ1 protein expression is localized specifically to ciliated cells (Blatt EN et al., 1999). Similar to the FoxJ1 gene, a combination of beta-tubulin and FoxJ1 markers associates with ciliated cell differentiation (Lazard DS et al., 2009, LeSimple P et al., 2007, Rowe RK et al., 2004).

The use of keratins 5 and 14 as markers for basal cell regeneration during wound repair has been reported previously (Dakir EL et al., 2008, Patel GK et al., 2006) and the expression of keratin 14 gene has been reported to be time-dependent, decreasing in extent during the process of wound repair (Lazard DS et al., 2009). A study by Patel GK et al., (2006) revealed the presence of six new keratinocyte subpopulations in the epidermis during acute wound healing (2–8 days), as defined by their differing keratin expression, and amongst those keratinocyte genes, keratin 14 (KRT14) and KRT14 were expressed in basal and immediate suprabasal cells of the epidermis adjacent to the wound edge (Patel GK et al., 2006). These basal keratinocytes are fundamental for epidermal integrity and migrate over the wound site. A study by Liu JY et al., (1994) during epithelial regeneration in tracheal grafts suggested that keratin 14 appeared before markers of highly differentiated secretory or ciliated cells, and were instead associated with cells which were poorly differentiated and highly proliferative (Liu JY et al., 1994). In addition, Shimizu T et al., (1994) reported that following a focal denuding wound in mechanically injured rat trachea, cells from the adjacent non-wounded epithelium flattened and migrated into the wounded site. These populations of cells were reported to dedifferentiate to become a population of poorly differentiated (PD) cells which expressed keratin 14 and Griffonia simplicifolia I-isolectin B4 (GSI-B4) lectin binding sites, which are specific for basal cells in normal epithelium (Shimizu T et al., 1994). Thus, these findings suggest that dedifferentiated cells from adjacent sites adopt a basal cell-like

phenotype and play a role in migrating towards and replacing denuded areas in response to injury and it is believed that this process is one of the fundamental mechanisms of the repair process following injury.

With these literature-based observations and findings to hand, and with the overarching objective in place of developing an understanding of the mechanisms that underlay the different stages of wound repair in the airway epithelium in sheep, it was clear that specific needs could be identified that would be key to achieving this goal. Whilst protein markers would provide the means to trace the behaviour of individual cell types and their involvement in this process, in the majority of instances such markers are not yet available for use in sheep. Molecular characterisation of this process, on the other hand, is to some extent a more realisable current objective as many of the genes involved have been characterised in the sheep and/or are highly conserved between species. As such the opportunity exists to identify at least some of the likely key, or candidate, genes involved and generate the primers necessary to follow the expression of such in the repair process. With this in mind, we designed preliminary experiments in which we elected to study and quantify the mRNA expression level for five selected genes, namely beta-catenin, beta-tubulin, keratins 5 and 14, and FoxJ1 in the ovine airway in response to bronchial brushing by using reverse-transcriptase polymerase chain reaction (RT-PCR). Our expectation was that findings in this regard would lead to a basic understanding of the way in which expression of those genes would relate to the different stages of epithelial repair.

3.2 OBJECTIVES

1. To develop reverse transcription polymerase chain reaction (RT-PCR) assays for the ovine beta-catenin, keratins 5 and 14, beta-tubulin and FoxJ1
2. To generate ovine sequences for beta-catenin, beta-tubulin, keratins 5 and 14 and FoxJ1 genes
3. To study and quantify the expression patterns for these genes at various time points following injury

3.3 MATERIALS AND METHODS

3.3.1 Experimental design

The bronchial brushing procedure was performed at single predefined sites within the major airways of anaesthetised sheep (n = 4), at selected time points (1, 3, 7, 14, 21 and 28 days prior to post mortem; Protocol A). However each sheep was sampled on only three or four different occasions (Table 3.1 and Figure 3.1): Sheep 2683 (days 1, 3, 7 and 14), Sheep 2676 (days 1, 3, 21 and 28) and Sheep 8038 and 8039 (days 1, 3 and 7). Naïve sites from each sheep were assigned as controls. A comparison of gene expression in mRNA prepared from brushed and unbrushed regions (Figure 3.1) of the same airway was performed by RT-PCR. The protocols involved in anaesthesia (Chapter 2; Section 2.3.3), the bronchial brushing (Chapter 2; 2.3.4) and necropsy (Chapter 2; 2.3.5) were followed.

Sheep	Time points (days prior to necropsy)											
	0 (naive)	**1**		**3**		**7**		**14**		**21**		**28**	
2683	RCD	RVD1	LCD	RVD2	LC	RC	LVD2	RA	LVD1	-		-	
2676	RCD	LCD	RVD1	LC	RVD2	-		-		RC	LVD2	RA	LVD1
8038	RVD1	RC		LC		RVD2		-		-		-	
8039	RC	RVD1		LC		LVD1		-		-		-	

Table 3.1: The sheep, the location of the subsegmental airways subjected to bronchial brushing and the various time points studied prior to necropsy
(RA = right apical; RC = right cranial; LC = left cranial; RCD = right cranial diaphragmatic; LCD = left cranial diaphragmatic; RVD = right ventral diaphragmatic; LVD = left ventral diaphragmatic)

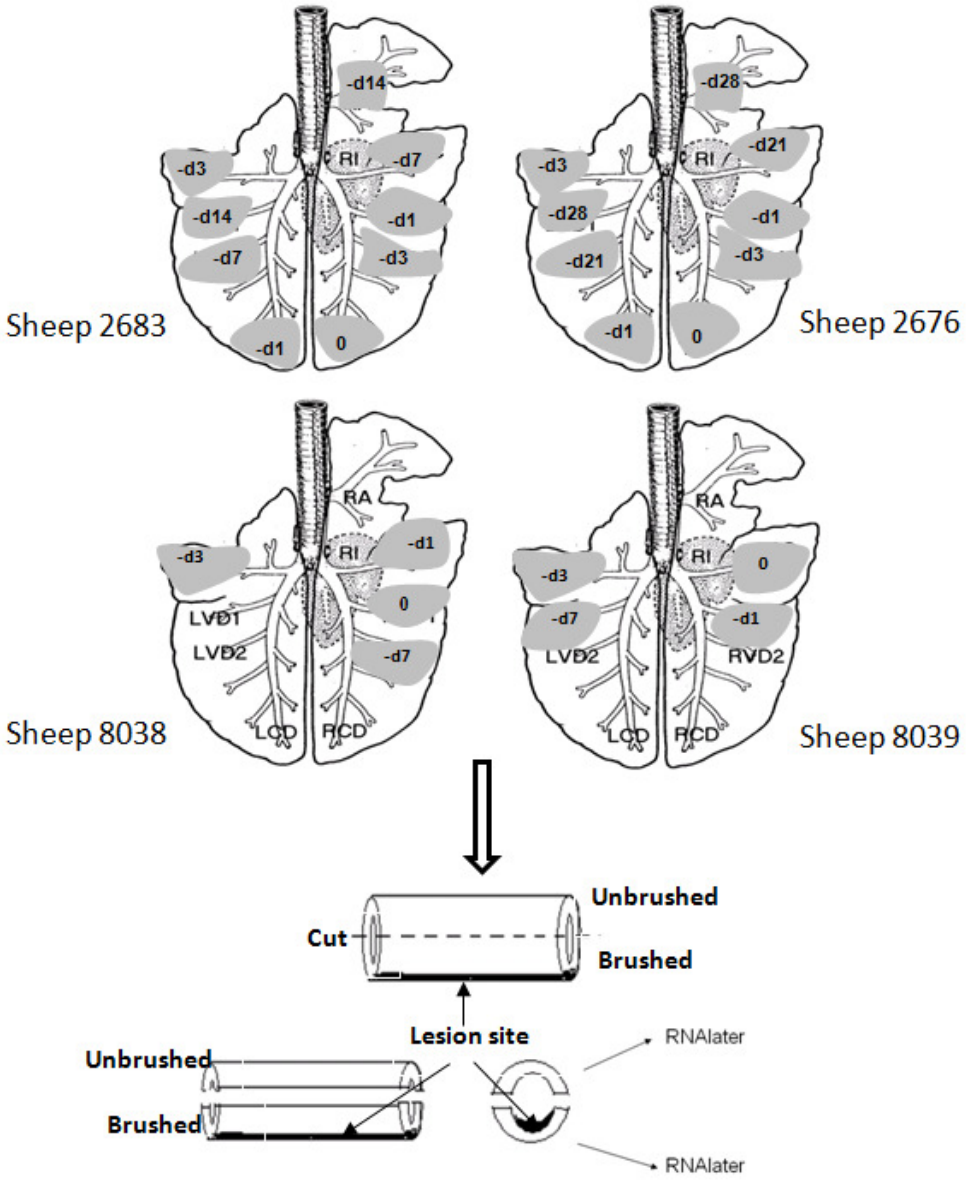


Figure 3.1: Schematic of the airway structures and the selected segments subjected to bronchial brushing at each specific time point

The airway segments were brushed at specific time points and during necropsy examination the airway segments were identified, dissected away from surrounding parenchyma and cut into two separate sites relating to the areas that were brushed, and the opposing areas that were not. These dissected pieces were preserved in RNAlater. The negative (-) symbol before each time point in the schematic indicates 'prior to post mortem'. (* D = Day, 0 = Naive)

3.3.2 Tissue sampling

Brushed areas were carefully identified during the necropsy examination and the airway segments containing the sites under consideration were removed intact from the surrounding parenchyma. Brushed and naïve segments were identified, and for brushed segments, the airway was divided into brushed and unbrushed portions (Figure 3.1). Samples from both areas were collected into RNAlater (Ambion, UK) and stored at -20^{o}C for subsequent RNA extraction.

3.3.3 RNA extraction

RNA was prepared using the Qiagen Allprep Mini kits (Qiagen, UK). For each sample thirty milligrams of lung tissue was weighed into a FastRNA ProGreen matrix tube (MP Biomedicals, UK) containing 0.7 ml RLT buffer with β-mercaptoethanol (1 ml RLT buffer:10 µl β-mercaptoethanol). The matrix tube was processed in a FastPrep® Instrument (Bio 101; Thermo Electron Corp., UK) for 40 seconds at a speed setting of 6.0 (followed the procedures as recommended by the FastRNA Pro Green Kit (MP Biomedical): Instruction manual). The upper phase was transferred to a Qiashredder column (Qiagen, UK) and was centrifuged at 13,000 rpm for 2 minutes. The supernatant was transferred into an AllPrep Column and centrifuged at 10,000 rpm for 30 seconds at room temperature.

One volume of 70% ethanol was added to the flow-through solution, mixed and transferred to an RNeasy spin column then centrifuged at 10,000 rpm for 15 seconds. The flow-through was discarded and the column washed with 350 µl of RW1 buffer. The column was then centrifuged for 15 seconds (10,000 rpm at room temperature). After the second flow-through was discarded, 80 µl DNAse (Qiagen, UK) solution (10 µl of DNAse I stock solution: 70 µl buffer RDD) was added directly to the RNeasy spin column membrane and incubated at room temperature for 30 minutes. After incubation, 350 µl RWI was added to the spin column and centrifuged at 10,000 rpm for 15 seconds.

The column was transferred into a new collection tube and washed with 450 µl RPE/Ethanol buffer (1 RPE: 4 Ethanol). The column was centrifuged at 10,000 rpm

for 15 seconds and the flow-through was discarded. The washing step with RPE/Ethanol was repeated; the column was centrifuged at 10,000 rpm for 2 minutes and again the flow-through was discarded. The column was placed into a 1.5 mL eppendorf tube (Axygen Scientific, USA) and 50 µl of nuclease-free water was added to the column's membrane; the column was then centrifuged at 10,000 rpm for 1 minute to elute the RNA.

A 1:10 dilution of RNA sample was quantified using a Spectrophotometer at 260 nm, and purity was assessed by determining the OD ratio 260/280 nm.

3.3.4 Reverse transcription of mRNA into cDNA

A reverse transcription (RT) of 1 µg total RNA was set up on ice and cDNA copies were generated using a Perkin Elmer (Perkin-Elmer, USA) thermocycler. Briefly, 1µg of total RNA was denatured at 70^{o}C for 10 minutes and immediately chilled in ice for 5 minutes. The reaction mixture (Promega, UK) consisted of 4µl of $MgCl_2$ (25mM), 2µl of 10x Reverse Transcription Buffer, 2 µl of deoxynucleotide triphosphate (dNTP) mixture (10mM), 0.5µl Recombinant RNAsin Ribonuclease Inhibitor, 0.5µl of 15U avian myeloblastoma virus (AMV) reverse transcriptase, 0.5µl Oligo(dT)15 primer and 9.5µl of nuclease free water added to give a final volume of 20µl. The reaction was incubated for 1 hour at 42^{o}C, followed by deactivation of the enzyme at 95^{o}C for 5 minutes. The RT mixture was then chilled on ice for 10 minutes and subsequently stored at -20^{o}C for further application. The control for RNA was set up by replacing the RNA template with RNA-free water and the same protocols were applied.

3.3.5 Oligonucleotide primers

The human sequences for beta-catenin, beta-tubulin, keratins 5 and 14, and FoxJ1 formed the basis for searching for available ovine or bovine sequences corresponding to those genes in the genomic sequence database (GenBank). The obtained bovine and ovine (or ovine EST) sequences for genes of beta-catenin, beta-tubulin, keratins 5 and 14 and FoxJ1 were aligned against available mammalian gene sequences using the BLAST algorithm through NCBI (http://blast.ncbi.nlm.nih.gov/Blast), Ensembl

(http://trace.ensembl.org/cgi-bin/tracesearch), European Bioinformatic Institute (EMBL-EBI) (http://www.ebi.ac.uk/blast/) and DFCI Sheep Gene Index databases, to identify areas of conserved sequence that might prove valuable in the context of primer design for the present application.

Primers for beta-tubulin and keratin 5 genes were designed based on the bovine sequence following inter-species sequence alignments. For keratin 5, primers were designed over conserved regions following alignment of bovine (NM_001008663), human (NM_000424) and mouse (NM_027011) sequences to give a 493bp PCR product. For the beta-tubulin gene, primers were designed based on conserved sequence between bovine (NM_001046549) and human (NM_178014) to amplify 503bp of PCR product.

A bovine expressed sequence tag (EST) (TC3860) for the keratin 14 gene was obtained from the DFCI database. ClustalW2 (EMBL-EBI) sequence alignment analysis revealed that the TC3860 bovine EST sequence was 97% similar to the bovine keratin 14 gene. This EST sequence was used to design the primers for the ovine keratin 14 gene based on conserved sequences identified from an alignment with horse, chicken, chimp, human and mouse sequence. Based on the bovine mRNA sequence, primers were expected to amplify a 407bp PCR product.

For the FoxJ1 gene, the primers were designed based on sheep genomic DNA sequence (DU231961: Ovis aries genomic clone CH243-427C5) identified from the EMBL-EBI database by blasting the human forkhead box J1 (FoxJ1) mRNA sequence (NM_001454.2). In order to check that this sequence was really coding for FoxJ1 gene homologue, the blast analysis was completed with the ovine sequence using the NCBI reference mRNA sequence (refseq-rna) database, and the results showed that 99% of this sequence is similar to the bovine Forkhead box protein J1 (Forkhead-related protein FKHL13, FoxJ1) mRNA (XM_582255.3) and 92% similar to that of human Forkhead box J1 (FoxJ1) mRNA (NM_001454.2). Since there was no available ovine mRNA sequence for FoxJ1 gene, primers were designed based on conserved regions identified from sequence alignment analysis between bovine

mRNA and ovine genomic DNA. The primers were designed to amplify a predicted 210bp of ovine FoxJ1 mRNA sequence (Figure 3.2).

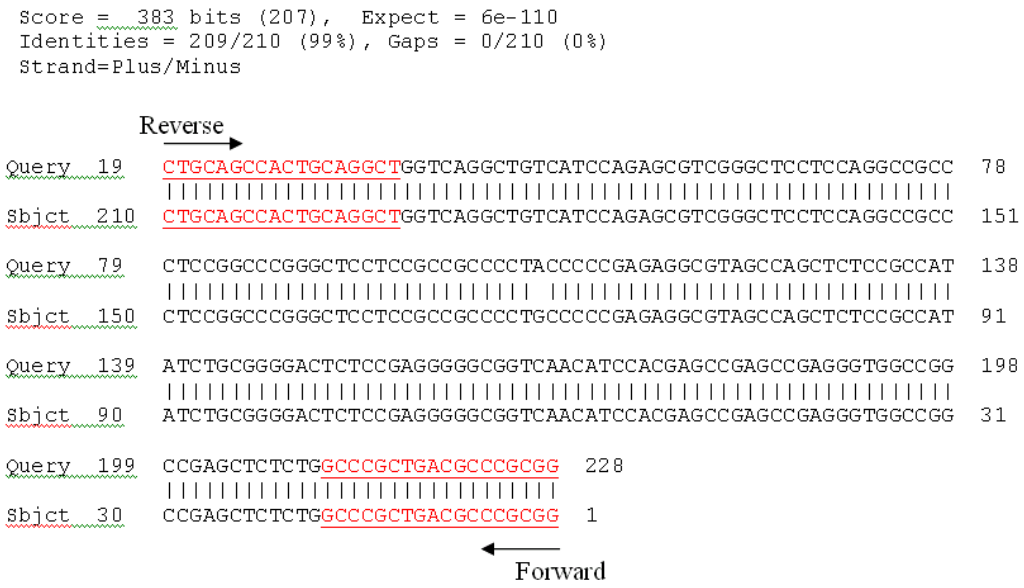


Figure 3.2: A sequence alignment between bovine mRNA and ovine genomic DNA to find the conserved areas appropriate for primer design. The primers were designed to amplify a predicted 210bp of ovine FoxJ1 mRNA sequence. These two sources of sequences are 99% similar.

Primers for the ovine beta-catenin (FE032823) gene were designed based on an ovine mRNA sequence obtained from the EMBL-EBI database by blast analysis with human beta-catenin. This sequence was aligned with that of bovine (NM_001076141), human (NM_001098210) and mouse (NM_007614) beta-catenin mRNA sequences, and the primers were designed at conserved regions. The primers were predicted to amplify a 407bp PCR product.

Primers were designed by using DNAsis Max v2.0 (Micobio Inc, USA). All designed primers were again checked for specificity using BLAST algorithm through NCBI http://blast.ncbi.nlm.nih.gov/Blast. Primers were purchased from Eurofin (MWG Operon, UK).

ATPase, a constitutively expressed 'housekeeping' gene, was used as a control to assess the quality of the RT product and also as a reference for quantification of

mRNA (Craig NM et al., 2007). The list of the primers used in this study is shown in (Table 3.2)

3.3.6 PCR amplification of sheep cDNA

Beta-tubulin, keratin 5 and beta-catenin genes were each amplified in a duplex PCR with ATPase whilst keratin 14 and FoxJ1 were run individually since the high annealing temperature required for their respective primers were incompatible with the ATPase.

PCR reactions were set up on ice using GoTaq Flexi DNA polymerase (Promega, UK) consisting of a 5x GoTaq Flexi Buffer and 25 mM $MgCl_2$ solution. PCR reactions, which were contained within thin-walled 0.2 ml PCR tubes (Axygen Scientific, USA), consisted of 2µl cDNA template, 5µl of 5x GoTaq Flexi Buffer, 4µl dNTPs (10mM) (Bioline, UK), 2µl $MgCl_2$ solution (25mM), 1µl of each forward and reverse primers (10pmol) (for duplex PCR, the 1µl of each forward and reverse primers for ATPase primer was added), 0.125µl of GoTaq Flexi DNA Polymerase (5u/µl) and 4.875µl nuclease free water added in a total volume of 20µl reactions.

GENES	ANNEALING TEMP (^{O}C)	CYCLES	SIZE (bp)	SEQUENCES
Beta-tubulin	57.3	28	503	Fwd: 5'-GTGGCAAATATGTCCCTCGT-3', Rev:5'-TCCATAGGTTGGTGTGGTCA-3'
Beta-catenin	60.0	28	407	Fwd: 5'-GCTGCAGGGGTCCTCTGTGA-3', Rev:5'-GGATAGTCAGCACCGGGGTGG-3'
Keratin 5	59.4	28	493	Fwd: 5'-CACCAAGACTGTGAGGCAGA-3', Rev:5'-AGTCTGCTGCAGCTCCTCAT-3'
Keratin 14	64.0	28	405	Fwd: 5'-GACATCAACGGCCTGCGCAG-3', Rev:5'-TGCTGAGCTGGGACTGCAG-3'
FoxJ1	65.0	35	210	Fwd: 5'-CCGCGGGCGTCAGCGGGC-3', Rev:5'-CTGCAGCCACTGCAGGCT-3'
ATPase	60*	28*	167	Fwd: 5'-GCTGACTTGGTCATCTGC-3', Rev: 5'-CAGGTAGGTTTGAGGGGATAC-3'

Table 3.2: The list of primer sequences and PCR conditions for amplifying the candidate genes in this study
The expected size of the PCR products is based on the length of both forward (Fwd) and reverse (Rev) primers expected to be annealed on the target regions. (*) the annealing temperature and PCR cycles varied according to the methods (using either singleplex or duplex PCR techniques).

The Techne Gradient (Techne, USA) and/or Perkin-Elmer GeneAmp PCR system 2400 (Perkin-Elmer, USA) PCR machines were used to heat the samples at 94^{o}C for 5 minutes for the initial denaturation. The PCR cycles were initiated with denaturation of the cDNA template at 94^{o}C for 30 seconds; the primers were annealed at the specific annealing temperatures for each gene (as shown in Table 3.2) for 30 seconds followed by the extension step at 72^{o}C for 30 seconds. The final extension step was completed at 72^{o}C for 7 minutes to make sure the target regions of the genes had been transcribed completely. For a duplex PCR, the PCR conditions for the ATPase gene were optimised and compatible with the PCR conditions of the genes of interest; however, for a semi-quantitative analysis of the FoxJ1 and keratin 14 genes, the ATPase gene was amplified alone at 35 and 28 cycles respectively using the following PCR conditions: initial denaturation for 2 minutes at 94^{o}C, followed by denaturation at 94^{o}C for 40 seconds, annealing at 60^{o}C for 40 seconds, extension at 72^{o}C for 2 minutes, followed by the final extension at 72^{o}C for 7 minutes.

A 10 µl aliquot of each PCR product was loaded into the well of a 1.3% TBE-agarose gel (Invitrogen, UK) for electrophoresis to check product size against 1kb DNA ladder (Promega, UK). Gels were stained with GelRed (Biosystem, UK) (10 µl/100 ml) and photographed using QuantityOne software (Bio-Rad Laboratories, USA).

Following purification of the PCR product using a PCR Purification kit (Qiagen, UK) according to the manufacturer's protocols, the identity of the PCR products was assessed by DNA sequencing. PCR products were sequenced in both directions (using both forward and reverse primers) using di-deoxy chain termination cycle sequencing (Applied Biosystem, UK) at DBS Genomics, Durham University, UK.

3.3.7 Densitometry of PCR products

Following image collection, Kodak 1D Analysis software was used to determine the density of the PCR product bands. To compare the mRNA expression for each gene across samples, a semi-quantitative analysis was performed by normalising the band

density of the PCR product for the specific gene of interest to the corresponding ATPase PCR product signal.

3.3.8 Molecular cloning

3.3.8.1 Preparation of cDNA template

High fidelity PCR products for all genes were prepared using Pfu *Taq* polymerase (Promega, UK). The PCR reactions, consisting of 5 µl of 10x Pfu buffer, 2 µl of 10 mM of each primer, 2 µl of 10 mM dNTPs (Bioline, UK), 4 µl of cDNA template, 0.25 µl of Pfu *Taq* polymerase and nuclease free water, were prepared in a total volume of 50 µl. The PCR reactions were amplified using the Perkin-Elmer GeneAmp PCR system 2400 (Perkin-Elmer, USA). The cDNA templates were first denatured at 95^{o}C for 2 minutes, and the PCR cycles were initiated with denaturation of the cDNA template at 95^{o}C for 30 seconds; the primers were annealed at the specific annealing temperatures for each specific gene (Table 3.2) for 30 seconds, followed by the extension step at 72^{o}C for 1 minute. The final extension step was done at 72^{o}C for 7 minutes. The PCR products were purified using QIAquick PCR Purification kit (Qiagen, UK).

3.3.8.2 A+-tailing procedures

Since PCR products amplified with proof-reading (Pfu) *Taq* polymerase are blunt-ended, the A-tailing procedure was carried out in order to generate A-tailed overhang at both ends to facilitate TA-cloning. The reactions were prepared in a total volume of 10 µl, and reaction mixtures were as follows: 5 µl of purified cDNA template, 1 µl of 10 mM dATP (Promega, UK), 1 µl of GoTaq Buffer (Promega, UK), 1 µl of GoTaq Flexi DNA polymerase (Promega, UK) and 2 µl of nuclease-free water. The reactions were incubated at 70^{o}C for 30 minutes using the Perkin-Elmer GeneAmp PCR system 2400 (Perkin-Elmer, USA).

3.3.8.3 Ligation into pGEM-T Easy Vector

A-tailed product was cloned into pGEM-T Easy Vector (Promega, UK). Ligation reactions contained: 5µl of 2x Rapid Ligation Buffer, 1 µl of pGEM-T Easy Vector, 1µl of T4 DNA Ligase, 25 ng of A-tailed cDNA products, and 1 µl of nuclease-free

water to a total volume of 10 µl. The ligation reactions were incubated overnight at 4^{o}C in the Perkin-Elmer GeneAmp PCR system 2400 (Perkin-Elmer, USA) machine.

3.3.8.4 Transformation into *E.coli* JM109

The ligation mixtures were spun down and 2 µl of ligation reactions were transferred to the bottom of sterile 1.5 ml centrifuge tubes on ice. An *E.coli* JM109 High Efficiency Competent Cells (>10^{8} cfu/µg DNA) (Promega, UK) were removed from -80^{o}C storage and placed in an ice bath until just thawed (5 minutes). The cells were gently mixed by flicking the tubes. JM109 (25 µl) cells were transferred into the tubes containing the ligation reactions and placed on ice for 20 minutes. The reactions were heat shocked for 50 seconds in a 42^{o}C water bath and immediately chilled on ice for another 2 minutes. 950 µl of SOC medium (Invitrogen, UK) was added and tubes were shaken at 150 rpm for 1.5 hours. After incubation, the reactions were spun at 3,000rpm for 10 minutes and re-suspended in 200 µl of SOC medium. 100 µl of cells were plated in duplicates on LB agar (Invitrogen, UK) containing ampicilin (100 µg/mL), IPTG (100 mM) and X-gal (50 mg/mL), and the plates were incubated overnight at 37^{o}C. Control for DNA insert (provided by the manufacturer) was used and the same protocols performed, with the exception of the A-tailing procedure.

3.3.8.5 Screening for positive clones

The numbers of white/blue colonies were counted and recorded. Single white colonies from the sub culture plate were selected for screening by PCR (the same protocols used for original PCR amplification using Go*Taq* polymerase). PCR-positive sub-cultures were selected for plasmid DNA extraction using Wizard® Plus DNA Purification System (Promega, USA) after overnight incubation in LB Broth (Invitrogen, UK), and the resulting recombinant plasmid DNAs were sent for sequencing using T7 and SP6 primers.

3.4 RESULTS

3.4.1 Molecular cloning and sequence validation

In the present study, PCR-based assays were developed in order to detect the expression of ovine beta-tubulin, beta-catenin, keratin 5, keratin 14 and FoxJ1 genes in samples derived from the ovine airway. In order to obtain and validate the sequence of amplified cDNA for all five genes, the PCR products were cloned into pGEM-T Easy vector and transformed into bacterial *E. coli* clone JM109. The results of the sequencing analysis for each gene are discussed below.

The results for direct sequencing of PCR products prior to cloning demonstrated that the sequence for beta-catenin showed 98% similarity to bovine (NM_001076141.1) and 94% similarity to human (NM_001904.3) beta-catenin mRNA sequences. The sequence for beta-tubulin showed 99% similarity to bovine (NM_001046549.1) and 89% similarity to human (NM_178014.2) beta-tubulin mRNA sequence. The sequence for keratin 5 showed 98% similarity to bovine (NM_001008663.1) and 90% similarity to human (NM_000424.3) keratin 5 mRNA sequences and the sequence for FoxJ1 showed 99% similarity to bovine (XM_582255.3) and 92% similarity to human (NM_001454.3) FoxJ1 mRNA sequences. Lastly, the sequence for keratin 14 showed 82% similarity to human keratin 14 (NM_000526.4) and 91% similarity to bovine keratin 13 (NM_001130747.1) mRNA sequences.

3.4.1.1 Screen for positive clones

PCR products amplified from two different RNA samples were used for cloning in duplicate. PCR products amplified for beta-catenin were successfully cloned. (Figure 3.3A) shows the PCR products derived from screening the positive clones. Clones A, C, D and E were complete 407bp sequences. Of the 29 clones analyzed, only 16 gave a PCR product of the expected 503bp size for beta-tubulin. Five clones were sent for DNA sequencing (Figure 3.3B) and 3 clones (clones E, C and M) had a complete 503bp sequence for the beta-tubulin gene with correct 5' to 3' orientation. The PCR products for keratin 5 were amplified using normal *Taq* polymerase (Go*Taq* polymerase) since proof-reading *Taq* polymerase (Pfu) gave no product despite attempts to optimise the assay. The PCR products were amplified from two different

sources of airway tissue (LC and RVD1), and the A+-tailing procedure was not required. These PCR products were successfully cloned and, of the 40 white colonies observed, 39 were shown to have a positive clone for the KRT14 gene when screened by PCR using specific primers (Figure 3.3C). Three clones contained the complete 493bp length of sequence (clones A, B and C). PCR products for the keratin 14 gene were successfully amplified and cloned (Figure 3.3D). Out of 6 clones sent for sequencing, only 2 clones produced a satisfactory result but were cloned in the wrong orientation (clones J and Q). The PCR products of this gene were amplified using normal *Taq* polymerase (Go*Taq* polymerase) since proof-reading *Taq* polymerase (Pfu) gave no product despite attempts to optimise the assay. The PCR products for FoxJ1 were successfully cloned and, out of four plates of colonies, only seven white colonies were observed. Among those, 6 were positive when screened by PCR using specific primers (Figure 3.3E), and these were sent for DNA sequencing analysis. Only 2 clones were successfully sequenced as having the complete 210bp in length (clones C and D). Low quality of sequencing data was the primary factor in leading to rejection of clones.

3.4.1.2 Sequence analyses

The sequencing results generated from both primers (T7 and SP6) were used to perform a blast 2-Sequence alignment to get full sequences of the genes of interest. For beta-catenin, although all samples had a complete target sequence (407bp) only four clones were cloned in 5' to 3' orientation. Sequence alignment of those four clones (ClustalW) demonstrated that they were identical. When this sequence was blasted through the EMBL-EBI (www.ebi.ac.uk) database, the results demonstrated 100% similarity with unpublished ovine EST sequences (EE756183, EE754439, FE022986, FE032823 and EE782646), and 98% similarity with bovine beta-catenin gene sequences (BC119949 and BT030683). Figure 3.4A shows the sequence alignment between clones which was performed using ClustalW on the EMBL-EBI database.

For beta-tubulin, sequence alignment (ClustalW) analysis of the 3 clones showed that these clones share a 100% similarity (Figure 3.4B). Blast analysis was performed

using EMBL-EBI and NCBI databases, and the results revealed that the cloned sequences are 99% homologous with the unclassified ovine ESTs (EE798210, EE765525, EE747478, DY496913, EE774297) and 98% similar to bovine beta-tubulin (TUBB) mRNA (BT030522 and NM_001046549.1 (NCBI)) and bovine beta-tubulin 2A mRNA (BC105401). Furthermore, these sequences are 90% similar to that of a human beta-tubulin mRNA sequence (NM_178014.2).

For keratin 5, seven clones were selected for full-length DNA sequencing and a blast 2-Sequence analysis was performed. Three clones had sequence of the anticipated 493bp size with the primer sequences at either end (Figure 3.4C) and cloned in correct orientation. When the sequence of all three clones were blasted through the NCBI database, the sequence was found to be 98% similar to that of the bovine keratin 5 mRNA sequence (NM_001008663.1) and 91% similar to the keratin 5 mRNA sequence for humans (NM_000424.3) and mice (NM_027011.2). When this ovine sequence was blasted through the EMBL-EBI database, a sequence was identified with a 100% similarity with unclassified Ovis aries (EE838877) EST sequence generated from wool follicle mRNA. Analysis using the NCBI database demonstrated that this EST sequence produced the same results as the 493bp clone. Therefore, it is likely that cloned sequence is the ovine keratin 5 mRNA sequence.

For keratin 14, the clones were not identical (79% identity) and both clones were 403bp rather than the 405bp predicted from the bovine sequence (TC3860) - (Figure 3.4D - clone J). Since the primers were designed based on the bovine EST sequence for the KRT14 gene (TC3860), this sequence might possibly reflect a species difference. The blast result for the clone KRT14-J, using the NCBI database revealed that this sequence demonstrated 98% similarity with bovine and human keratin 17 mRNA (NM_001105322.1 and NM_000422.2 respectively), 92% similarity with human keratin 14 mRNA (NM_000526.4) and 90% similarity with mouse keratin 14 mRNA sequence (NM_016958.1). In addition, when the result of this sequence was blasted through the DFCI sheep EST database, these results showed that 98% of this sequence was similar to bovine EST for keratin 17 gene sequence (TC2398 and

TC2943), 95% similar to bovine EST keratin 14 gene sequence (TC3860) and 81% similar to an ovine EST Keratin 15 gene sequence (TC6677).

However, analysis of the KRT14-Q clone sequence through the NCBI database revealed that this sequence demonstrated 100% similarity with K13 of Pan troglodytes (XM_511487), 91% similarity with bovine K15 (NM_001130747.1) and only 82% similarity with that of human KRT14 (NM_000526.4). Analysis using the DFCI database revealed that this sequence demonstrated 91% similarity with ovine K15 (TC6677), 80% similarity with bovine K17 (TC2398) and 80% similarity with bovine KRT14 (TC3860). It was therefore considered unlikely that the clone KRT14-Q encoded the KRT14 gene. Further analysis is required to confirm if clone KRT14-J genuinely codes for the keratin 14 gene.

For FoxJ1, sequence analysis showed that the two clones (both 210bp in length) were identical. A blast 2 Sequence analysis indicated that these clones shared similar sequence to the expected FoxJ1 gene (Figure 3.4E). In order to confirm that the sequences of these two clones belong to the FoxJ1 gene, a Blast analysis using the NCBI and EMBL-EBI databases was performed. Analysis of these sequences demonstrated a similarity of 99% with bovine Forkhead box protein J1 (Forkhead-related protein FKHL13) (Hepatocyte nuclear factor 3 forkhead homolog 4) (HNFH-4) (FoxJ1), mRNA sequence (XM_582255.3), and 92% similarity with human forkhead box J1 (FoxJ1), mRNA sequence (NM_001454.2).

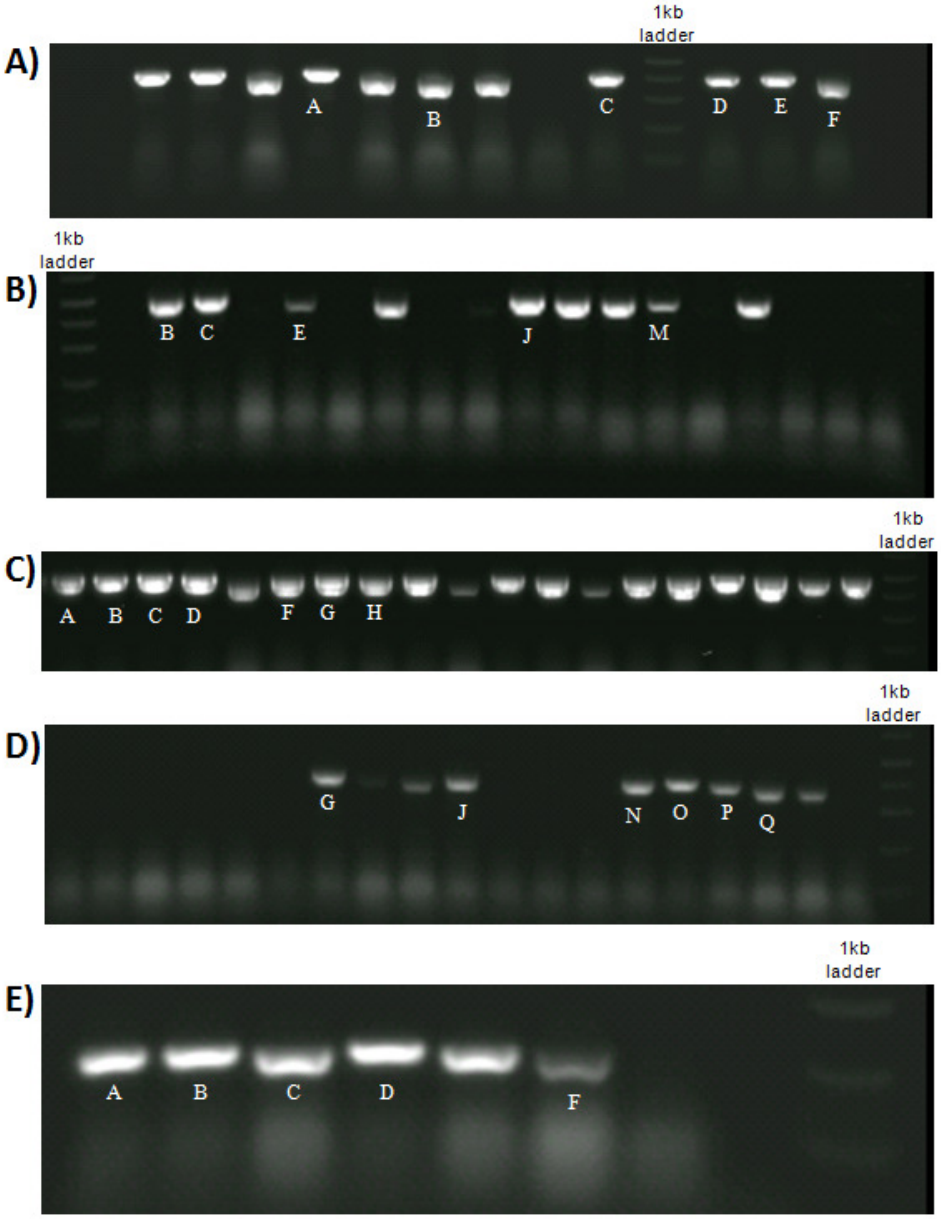


Figure 3.3: A screening for positive clones by PCR
The PCR products of the expected positive clones for the beta-catenin, beta-tubulin, KRT14, KRT14 and FoxJ1 genes were screened by PCR using specific primers. The white colonies were used as PCR template and different sizes of the PCR products observed indicate that some of the clones might represent the full sequence of the genes of interest. (the lane characters indicate the samples selected for DNA sequencing); A) beta-catenin, B) beta-tubulin, C) Keratin 5, D) Keratin 14, and E) FoxJ1

(A)

```
CloneD      GCTGCAGGGGTCCTCTGTGAACTTGCTCAGGACAAGGAAGCTGCAGAAGCTATTGAAGCC 60
CloneE      GCTGCAGGGGTCCTCTGTGAACTTGCTCAGGACAAGGAAGCTGCAGAAGCTATTGAAGCC 60
CloneC      GCTGCAGGGGTCCTCTGTGAACTTGCTCAGGACAAGGAAGCTGCAGAAGCTATTGAAGCC 60
CloneA      GCTGCAGGGGTCCTCTGTGAACTTGCTCAGGACAAGGAAGCTGCAGAAGCTATTGAAGCC 60
            ************************************************************
CloneD      GAGGGAGCCACAGCTCCTCTGACAGAACTACTTCATTCTAGGAATGAGGGTGTGGCAACA 120
CloneE      GAGGGAGCCACAGCTCCTCTGACAGAACTACTTCATTCTAGGAATGAGGGTGTGGCAACA 120
CloneC      GAGGGAGCCACAGCTCCTCTGACAGAACTACTTCATTCTAGGAATGAGGGTGTGGCAACA 120
CloneA      GAGGGAGCCACAGCTCCTCTGACAGAACTACTTCATTCTAGGAATGAGGGTGTGGCAACA 120
            ************************************************************
CloneD      TATGCAGCTGCTGTTTTGTTCCGAATGTCTGAGGACAAGCCACAGGATTATAAGAAACGG 180
CloneE      TATGCAGCTGCTGTTTTGTTCCGAATGTCTGAGGACAAGCCACAGGATTATAAGAAACGG 180
CloneC      TATGCAGCTGCTGTTTTGTTCCGAATGTCTGAGGACAAGCCACAGGATTATAAGAAACGG 180
CloneA      TATGCAGCTGCTGTTTTGTTCCGAATGTCTGAGGACAAGCCACAGGATTATAAGAAACGG 180
            ************************************************************
CloneD      CTTTCAGTTGAGCTGACCAGTTCTCTCTTCAGAACGGAGCCAATGGCTTGGAATGAGACT 240
CloneE      CTTTCAGTTGAGCTGACCAGTTCTCTCTTCAGAACGGAGCCAATGGCTTGGAATGAGACT 240
CloneC      CTTTCAGTTGAGCTGACCAGTTCTCTCTTCAGAACGGAGCCAATGGCTTGGAATGAGACT 240
CloneA      CTTTCAGTTGAGCTGACCAGTTCTCTCTTCAGAACGGAGCCAATGGCTTGGAATGAGACT 240
            ************************************************************
CloneD      GCTGATCTTGGACTTGATATTGGTGCCCAGGGAGAACCTCTTGGATATCGCCAGGATGAT 300
CloneE      GCTGATCTTGGACTTGATATTGGTGCCCAGGGAGAACCTCTTGGATATCGCCAGGATGAT 300
CloneC      GCTGATCTTGGACTTGATATTGGTGCCCAGGGAGAACCTCTTGGATATCGCCAGGATGAT 300
CloneA      GCTGATCTTGGACTTGATATTGGTGCCCAGGGAGAACCTCTTGGATATCGCCAGGATGAT 300
            ************************************************************
CloneD      CCCAGCTATCGTTCTTTTCACTCCGGTGGGTACGGCCAGGACGCCTTGGGGATGGACCCC 360
CloneE      CCCAGCTATCGTTCTTTTCACTCCGGTGGGTACGGCCAGGACGCCTTGGGGATGGACCCC 360
CloneC      CCCAGCTATCGTTCTTTTCACTCCGGTGGGTACGGCCAGGACGCCTTGGGGATGGACCCC 360
CloneA      CCCAGCTATCGTTCTTTTCACTCCGGTGGGTACGGCCAGGACGCCTTGGGGATGGACCCC 360
            ************************************************************
CloneD      ATGATGGAGCATGAGATGGGCGGCCACCACCCCGGTGCTGACTATCC 407
CloneE      ATGATGGAGCATGAGATGGGCGGCCACCACCCCGGTGCTGACTATCC 407
CloneC      ATGATGGAGCATGAGATGGGCGGCCACCACCCCGGTGCTGACTATCC 407
CloneA      ATGATGGAGCATGAGATGGGCGGCCACCACCCCGGTGCTGACTATCC 407
            ***********************************************
```

(B)

```
CloneM      GTGGCAAATATGTCCCTCGTGCTATCTTGGTGGATCTAGAACCCGGAACCATGGACTCTG 60
CloneC      GTGGCAAATATGTCCCTCGTGCTATCTTGGTGGATCTAGAACCCGGAACCATGGACTCTG 60
CloneE      GTGGCAAATATGTCCCTCGTGCTATCTTGGTGGATCTAGAACCCGGAACCATGGACTCTG 60
            ************************************************************

CloneM      TTCGCTCAGGTCCTTTTGGTCAGATCTTCAGACCAGACAACTTTGTTTTTGGTCAGTCCG 120
CloneC      TTCGCTCAGGTCCTTTTGGTCAGATCTTCAGACCAGACAACTTTGTTTTTGGTCAGTCCG 120
CloneE      TTCGCTCAGGTCCTTTTGGTCAGATCTTCAGACCAGACAACTTTGTTTTTGGTCAGTCCG 120
            ************************************************************

CloneM      GGGCAGGCAACAACTGGGCCAAGGGCCACTATACAGAGGGAGCCGAGCTGGTTGACTCGG 180
CloneC      GGGCAGGCAACAACTGGGCCAAGGGCCACTATACAGAGGGAGCCGAGCTGGTTGACTCGG 180
CloneE      GGGCAGGCAACAACTGGGCCAAGGGCCACTATACAGAGGGAGCCGAGCTGGTTGACTCGG 180
            ************************************************************

CloneM      TCCTGGATGTGGTTCGGAAGGAGGCCGAGAGCTGTGACTGCCTGCAGGGCTTCCAGCTGA 240
CloneC      TCCTGGATGTGGTTCGGAAGGAGGCCGAGAGCTGTGACTGCCTGCAGGGCTTCCAGCTGA 240
CloneE      TCCTGGATGTGGTTCGGAAGGAGGCCGAGAGCTGTGACTGCCTGCAGGGCTTCCAGCTGA 240
            ************************************************************

CloneM      CCCACTCACTGGGTGGGGGTACTGGCTCCGGAATGGGCACCTTGCTCATTAGCAAGATCC 300
CloneC      CCCACTCACTGGGTGGGGGTACTGGCTCCGGAATGGGCACCTTGCTCATTAGCAAGATCC 300
CloneE      CCCACTCACTGGGTGGGGGTACTGGCTCCGGAATGGGCACCTTGCTCATTAGCAAGATCC 300
            ************************************************************

CloneM      GTGAAGAGTATCCTGACCGAATCATGAACACCTTCAGCGTGGTGCCCTCACCCAAAGTGT 360
CloneC      GTGAAGAGTATCCTGACCGAATCATGAACACCTTCAGCGTGGTGCCCTCACCCAAAGTGT 360
CloneE      GTGAAGAGTATCCTGACCGAATCATGAACACCTTCAGCGTGGTGCCCTCACCCAAAGTGT 360
            ************************************************************

CloneM      CCGACACTGTGGTTGAGCCCTACAATGCCACCCTCTCCGTCCATCAGCTGGTGGAGAACA 420
CloneC      CCGACACTGTGGTTGAGCCCTACAATGCCACCCTCTCCGTCCATCAGCTGGTGGAGAACA 420
CloneE      CCGACACTGTGGTTGAGCCCTACAATGCCACCCTCTCCGTCCATCAGCTGGTGGAGAACA 420
            ************************************************************

CloneM      CTGATGAAACCTACTGCATTGACAACGAGGCCCTTTACGACATATGTTTCCGCACCCTCA 480
CloneC      CTGATGAAACCTACTGCATTGACAACGAGGCCCTTTACGACATATGTTTCCGCACCCTCA 480
CloneE      CTGATGAAACCTACTGCATTGACAACGAGGCCCTTTACGACATATGTTTCCGCACCCTCA 480
            ************************************************************

CloneM      AGCTGACCACACCAACCTATGGA 503
CloneC      AGCTGACCACACCAACCTATGGA 503
CloneE      AGCTGACCACACCAACCTATGGA 503
            ***********************
```

(C)

```
CloneA          CACCAAGACTGTGAGGCAGAACCTGGAGCCTTTGTTCGAGCAGTACATCAACAACCTCAG 60
CloneC          CACCAAGACTGTGAGGCAGAACCTGGAGCCTTTGTTCGAGCAGTACATCAACAACCTCAG 60
CloneB          CACCAAGACTGTGAGGCAGAACCTGGAGCCTTTGTTCGAGCAGTACATCAACAACCTCAG 60
                ************************************************************

CloneA          GAGACAGCTGGATGGTATCCTTGGGGAGAGAGGCCGCCTGGACTCGGAGCTCAGGAACAT 120
CloneC          GAGACAGCTGGATGGTATCCTTGGGGAGAGAGGCCGCCTGGACTCGGAGCTCAGGAACAT 120
CloneB          GAGACAGCTGGATGGTATCCTTGGGGAGAGAGGCCGCCTGGACTCGGAGCTCAGGAACAT 120
                ************************************************************

CloneA          GCAGGACTTGGTGGAGGACTTCAAGAACAAGTATGAAGATGAAATCAACAAGCGTACCAC 180
CloneC          GCAGGACTTGGTGGAGGACTTCAAGAACAAGTATGAAGATGAAATCAACAAGCGTACCAC 180
CloneB          GCAGGACTTGGTGGAGGACTTCAAGAACAAGTATGAAGATGAAATCAACAAGCGTACCAC 180
                ************************************************************

CloneA          TGCCGAGAATGAGTTTGTGATGCTGAAGAAGGATGTGGACGCTGCCTACATGAACAAAGT 240
CloneC          TGCCGAGAATGAGTTTGTGATGCTGAAGAAGGATGTGGACGCTGCCTACATGAACAAAGT 240
CloneB          TGCCGAGAATGAGTTTGTGATGCTGAAGAAGGATGTGGACGCTGCCTACATGAACAAAGT 240
                ************************************************************

CloneA          GGAACTAGAGGCCAAGGTCGATGCACTGATGGACGAGATCAACTTCATGAAGATGTTCTT 300
CloneC          GGAACTAGAGGCCAAGGTCGATGCACTGATGGACGAGATCAACTTCATGAAGATGTTCTT 300
CloneB          GGAACTAGAGGCCAAGGTCGATGCACTGATGGACGAGATCAACTTCATGAAGATGTTCTT 300
                ************************************************************

CloneA          CGATGCGGAGCTGTCCCAGATGCAGACACATGTCTCAGACACATCCGTGGTCCTGTCCAT 360
CloneC          CGATGCGGAGCTGTCCCAGATGCAGACACATGTCTCAGACACATCCGTGGTCCTGTCCAT 360
CloneB          CGATGCGGAGCTGTCCCAGATGCAGACACATGTCTCAGACACATCCGTGGTCCTGTCCAT 360
                ************************************************************

CloneA          GGACAACAACCGCAGCCTGGACTTGGACAGCATCATTGCTGAGGTCAAGGCCCAGTATGA 420
CloneC          GGACAACAACCGCAGCCTGGACTTGGACAGCATCATTGCTGAGGTCAAGGCCCAGTATGA 420
CloneB          GGACAACAACCGCAGCCTGGACTTGGACAGCATCATTGCTGAGGTCAAGGCCCAGTATGA 420
                ************************************************************

CloneA          AGACATTGCCAACCGCAGCCGCACAGAAGCTGAGTCCTGGTACCAGACCAAGTATGAGGA 480
CloneC          AGACATTGCCAACCGCAGCCGCACAGAAGCTGAGTCCTGGTACCAGACCAAGTATGAGGA 480
CloneB          AGACATTGCCAACCGCAGCCGCACAGAAGCTGAGTCCTGGTACCAGACCAAGTATGAGGA 480
                ************************************************************

CloneA          GCTGCAGCAGACT 493
CloneC          GCTGCAGCAGACT 493
CloneB          GCTGCAGCAGACT 493
                *************
```

(D)

```
CloneJ          TGCTGAGCTGGGACTGCAGCTCGATCTCCAGGGCCTGCAAGGTGCGCCGGAGCTCCGAGA 60
CloneQ          TGCTGAGCTGGGACTGCAGCTCGATCTCCAGGCCCTGGAGCGTGCGCCTGAGCTCGGTGA 60
                ********************************* **** *  ******* ****** * **

CloneJ          TCTCGCTCTTGCCGCTCTGCACCAGCTCGCTGTTGGTGGCCACCTCGCGGTTCAGTTCCT 120
CloneQ          TCTCTGTCTTGCTGGTCTGGATCATGGCTGTGCTGGAAGACACCTCCTTGTTCAGCTCGG 120
                ****  ****** * **** * **   *  ** ***  * ******   ****** **

CloneJ          CCGTCTTGCTGAAGAACCAGTCCTCGGCATCCTTGCGGTTCTTCTCTGCCATCTTCTCAT 180
CloneQ          CGCTCTTGCTGTGGAACCATTCCTCGGCATCCCGGCGGTTCTTCTCTGCCATGGCCTCGT 180
                *  ********  ****** ************  ******************   *** *

CloneJ          ACTGGTCGCGCATCTCGTTCAGGATGCGGCTCAGGTCCACGCCGGGGGCGGCGTCCATCT 240
CloneQ          ACTGCTCCCTCATCTCTGCCAGCACGCGGGTCAGGTCAATGCCTGGTGCTGCGTCCATCT 240
                **** ** * ******   *** * **** ******* * *** ** ** **********

CloneJ          CCACGTTGATCTCGCCACCCACCTGGCCTCGAAGGGCTTTCATCTCCTCCTCGTGGTTCT 300
CloneQ          CCACGTTGACCTGGCCAGCCAGCTGGTTGCTGAACTCCTTCATCTCCTCCTCGTGGTTCT 300
                ********* ** **** *** ****   *  *   * **********************

CloneJ          TGCGGAGGTAGGCCAGCTCCTCCTTGAGGTTCTCGATCTGCATCTCCAGGTCAGCTCTGG 360
CloneQ          TCTTCAGGTAGGCCAGCTCCTCGTTCAGGCTCTCAATCTGCATCTCCAGATCAGTCTTGG 360
                *    ***************** ** *** **** ************** ****   ***

CloneJ          CCAGGGTCAGCTCGTCCAGCACCCTGCGCAGGCCGTTGATGTC 403
CloneQ          TCAGGGTCAGCTCGTCCAGCACCCTGCGCAGGCCGTTGATGTC 403
                 *******************************************
```

(E)

```
cloneC          CCGCGGGCGTCAGCGGGCCAGAGAGCTCGGCCGGCCACCCTCGGCTCGGCTCGTGGATGT 60
cloneD          CCGCGGGCGTCAGCGGGCCAGAGAGCTCGGCCGGCCACCCTCGGCTCGGCTCGTGGATGT 60
                ************************************************************

cloneC          TGACCGCCCCCTCGGAGAGTCCCCGCAGATATGGCGGAGAGCTGGCTACGCCTCTCGGGG 120
cloneD          TGACCGCCCCCTCGGAGAGTCCCCGCAGATATGGCGGAGAGCTGGCTACGCCTCTCGGGG 120
                ************************************************************

cloneC          GTAGGGGCGGCGGAGGAGCCCGGGCCGGAGGGCGGCCTGGAGGAGCCCGACGCTCTGGAT 180
cloneD          GTAGGGGCGGCGGAGGAGCCCGGGCCGGAGGGCGGCCTGGAGGAGCCCGACGCTCTGGAT 180
                ************************************************************

cloneC          GACAGCCTGACCAGCCTGCAGTGGCTGCAG 210
cloneD          GACAGCCTGACCAGCCTGCAGTGGCTGCAG 210
                ******************************
```

Figure 3.4: The sequence alignment analyses performed for beta-catenin, beta-tubulin, keratin 5, keratin 14 and FoxJ1 genes using ClustalW

The results show the results of sequence alignment analyses performed for the beta-catenin (A), beta-tubulin (B), keratin 5 (C), keratin 14 (D) and FoxJ1 (E) genes using ClustalW (EMBL-EBI). The comparisons were made between clones for each gene and showed 100% similarity over the complete sequence of the targeted genes except for keratin 14. In the latter instance the two clones only shared 80% identity. The underlined sequences show the sequences for the forward and reverse primers.

3.4.2 RT-PCR analysis of gene expression studies

The limited number of sheep used in the present study predicted an approach based on qualitative appraisal of the nature and extent of change in gene expression as a function of time after injury. Data are presented both in relation to individual values at each time point (wherein samples were derived from different segments), for each sheep, and also where the mean of the values at each time point was calculated – involving averaging of values derived from different segments and from different animals. Fold changes were calculated relating expression values in treated segments (from both brushed and unbrushed sites) to those obtained in naive segments (time point 0) for each individual sheep. Graphs were therefore plotted 1) based on the fold changes for brushed and unbrushed segments relative to the naive airways (left panel) and 2) based on the mean values (±SD) of absolute densitometry of genes of interest relative to absolute densitometry of ATPase gene for each time point (right panel) so as to compare overall expression patterns between brushed and unbrushed areas. Plotting the individual values (fold change) for each sample was deemed useful in assessing the extent of variation in expression between different sheep and duplicate segments within sheep.

Figure 3.5(a), Figure 3.6(a), Figure 3.7(a), Figure 3.8(a) and Figure 3.9(a) show the PCR products for the brushed, unbrushed and naive samples for the expression of the studied genes relative to ATPase. Figure 3.5(b), Figure 3.6(b), Figure 3.7(b), Figure 3.8(b) and Figure 3.9(b) show the expression values based on relative comparison between samples and ATPase gene plotted according to their fold change (left panel) and the mean values (right panel).

The graph (Figure 3.5(b); right panel) depicting the means and standard deviations of the expression relative to ATPase for brushed and unbrushed sites indicates that initially elevated levels of gene expression for beta-catenin on days 1 and 3 subsequently declined to broadly consistent levels of expression from day 7 onwards. Expression of beta-catenin within the brushed sites remained marginally more elevated than unbrushed sites at days 7, 14 and 28 post injury. Assessment of the individual fold-changes (Figure 3.5(b); left panel) contributing to these observations

indicates that there was considerable variation in expression both within and between sheep.

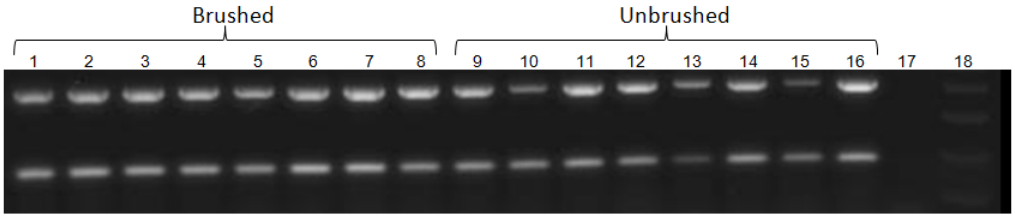


Figure 3.5(a): Agarose gel for a semi quantitative RT-PCR analysis for beta-catenin
Duplex RT-PCR products of beta-catenin and ATPase genes amplified from brushed (lanes 1-8) and unbrushed (lanes 9-16) samples at days 1 and 3 post injury, and a PCR negative control (lane 17).

Lane 1 & 9: Ovine 2683, segment LVD1, day 1

Lane 2 & 10: Ovine 2683, segment LCD, day 1

Lane 3 & 11: Ovine 2676, segment LCD, day 1

Lane 4 & 12: Ovine 2676, segment RVD1, day 1

Lane 5 & 13: Ovine 2683, segment RVD2, day 3

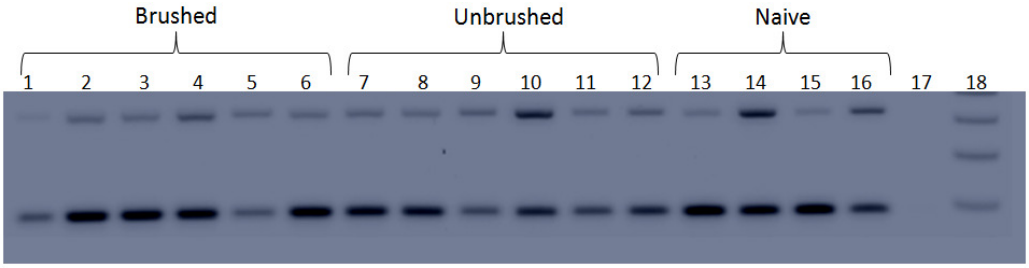
Lane 6 & 14: Ovine 2683, segment LC, day 3

Lane 7 & 15: Ovine 2676, segment LC, day 3

Lane 8 & 16: Ovine 2676, segment RVD2, day 3

Lane 17 : PCR negative control

Lane 18 : 1 kb DNA ladder



(...cont) **Agarose gel for a semi quantitative RT-PCR analysis for beta-catenin**
Duplex RT-PCR products of beta-catenin and ATPase genes amplified from brushed (lanes 1-6) and unbrushed (lanes 7-12) samples at days 1, 3 and 7 post injury, and a PCR negative control (lane 17).

Lane 1 & 7: Ovine 8038, segment RC, day 1

Lane 2 & 8: Ovine 8039, segment RVD1, day 1

Lane 3 & 9: Ovine 8038, segment LC, day 3

Lane 4 & 10: Ovine 8039, segment LC, day 3

Lane 5 & 11: Ovine 8038, segment RVD2, day 7

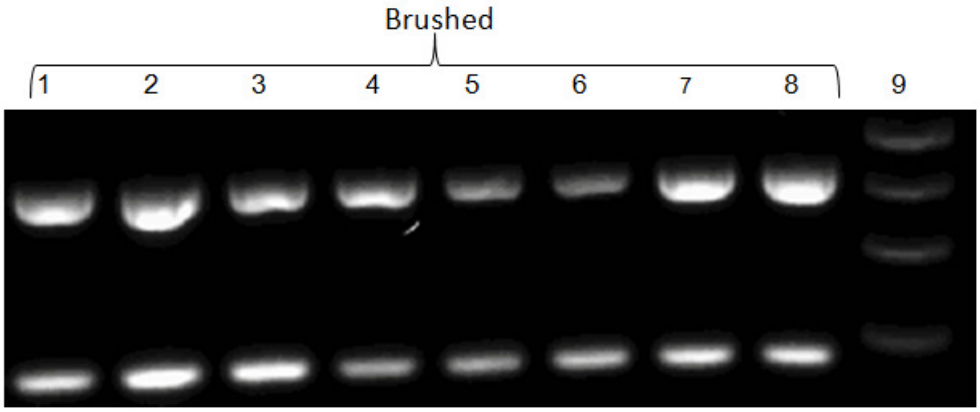
Lane 6 & 12: Ovine 8039, segment LVD1, day 7

Lane 13 & 14: Ovine 8038, segment RVD1, naive (duplicate)

Lane 15 & 16: Ovine 8039, segment RC, naive (duplicate)

Lane 17 : PCR negative control

Lane 18 : 1 kb DNA ladder



(*cont…*) Duplex RT-PCR products of beta-catenin and ATPase genes amplified from brushed samples at days 7, 14, 21 and 28 post injuries

Lane 1: Ovine 2683, segment LVD2, day 7

Lane 2: Ovine 2683, segment RC, day 7

Lane 3: Ovine 2683, segment LVD1, day 14

Lane 4: Ovine 2683, segment RA, day 14

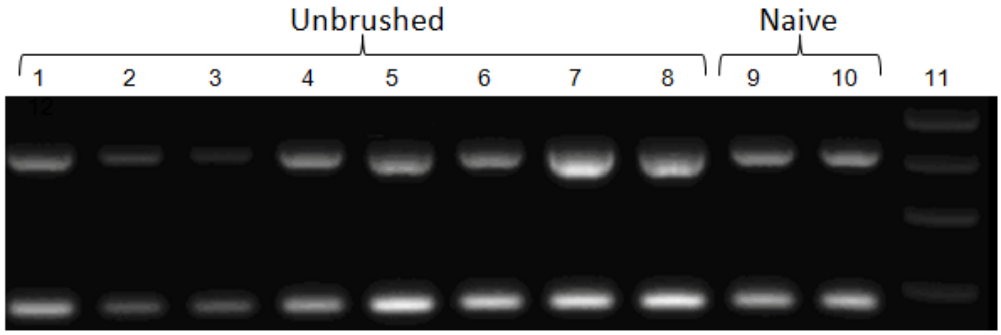
Lane 5: Ovine 2676, segment LVD2, day 21

Lane 6: Ovine 2676, segment RC, day 21

Lane 7: Ovine 2676, segment LVD1, day 28

Lane 8: Ovine 2676, segment RA, day 28

Lane 9: 1kb DNA ladder



(*cont…*) Duplex RT-PCR products of beta-catenin and ATPase genes amplified from unbrushed samples at days 7, 14, 21 and 28 post injuries

Lane 1: Ovine 2683, segment LVD2, day 7

Lane 2: Ovine 2683, segment RC, day 7

Lane 3: Ovine 2683, segment LVD1, day 14

Lane 4: Ovine 2683, segment RA, day 14

Lane 5: Ovine 2676, segment LVD2, day 21

Lane 6: Ovine 2676, segment RC, day 21

Lane 7: Ovine 2676, segment LVD1, day 28

Lane 8: Ovine 2676, segment RA, day 28

Lane 9: Ovine 2683, segment RCD, naïve (d14)

Lane 10: Ovine 2683, segment RCD, naïve (d28)

Lane 11: 1kb DNA ladder

Lane 12: PCR negative control

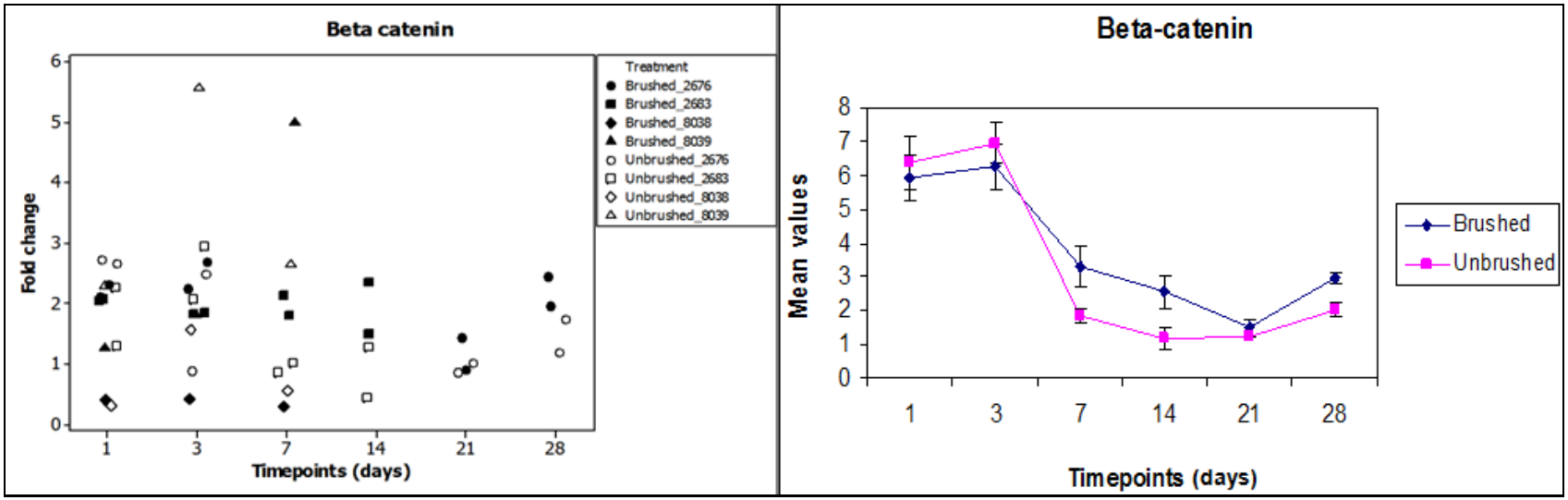


Figure 3.5 (b): Graphs depicting the expression patterns of beta-catenin between 1 and 28 days after injury. Graphs were plotted based on the fold change of the treated airways (brushed and unbrushed) relative to their respective naive airway (left panel) and based on the mean values (±SD). Expression of beta-catenin within the brushed sites remained marginally more elevated than unbrushed sites especially at days 7 and 14 post injury (right panel) whilst the individual fold changes plots show a considerable variation in expression of this gene (left panel).

The expression profile of keratin 5 in the brushed, relative to the unbrushed, area was up-regulated throughout the initial time points (Figure 3.6(b); right panel). The highest expression values observed in the brushed samples for this gene were 30-fold in excess of naïve site expression levels (Figure 3.6(b); left panel). The graph depicting averaged values relative to ATPase indicates an initial upregulation (days 1 and 3) of gene expression in the brushed sites (relative to unbrushed) that decreased through day 7 to levels broadly similar to those of ATPase by day 14 onwards for both brushed and unbrushed sites. However, as with beta-catenin there was evidence of variation within- and between-sheep (Figure 3.6(b); left panel).

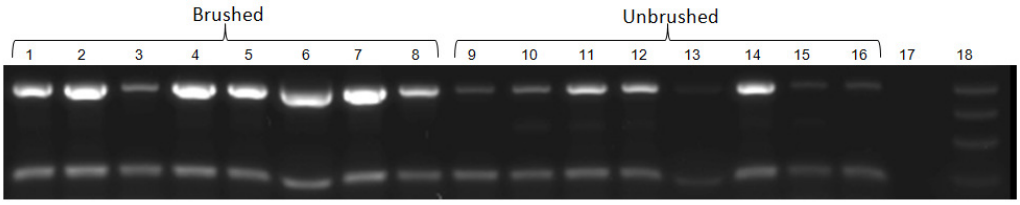


Figure 3.6 (a): Agarose gel for a semi quantitative of RT-PCR analysis for the keratin 5 gene
Duplex RT-PCR products of KRT5 and ATPase genes amplified from brushed (lanes 1-8) and unbrushed (lanes 9-16) samples at days 1 and 3 post injuries, and a PCR negative control (lane 17).

Lane 1 & 9: Ovine 2683, segment LVD1, day 1

Lane 2 & 10: Ovine 2683, segment LCD, day 1

Lane 3 & 11: Ovine 2676, segment LCD, day 1

Lane 4 & 12: Ovine 2676, segment RVD1, day 1

Lane 5 & 13: Ovine 2683, segment RVD2, day 3

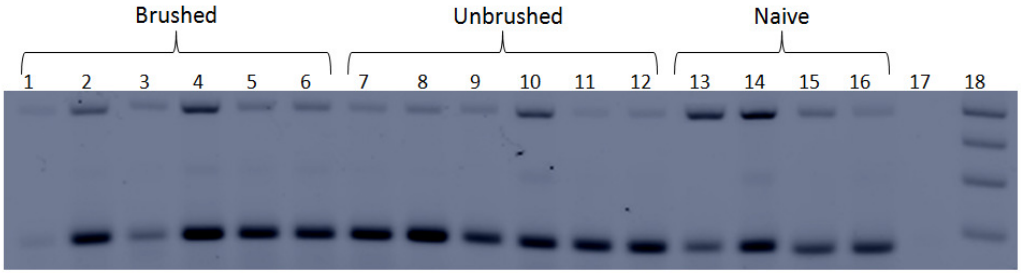
Lane 6 & 14: Ovine 2683, segment LC, day 3

Lane 7 & 15: Ovine 2676, segment LC, day 3

Lane 8 & 16: Ovine 2676, segment RVD2, day 3

Lane 17 : PCR negative control

Lane 18 : 1 kb DNA ladder



(...cont) **Agarose gel for a semi quantitative RT-PCR analysis for keratin 5**
Duplex RT-PCR products of KRT5 and ATPase genes amplified from brushed (lanes 1-6) and unbrushed (lanes 7-12) samples at days 1, 3 and 7 post injury, and a PCR negative control (lane 17).

Lane 1 & 7: Ovine 8038, segment RC, day 1

Lane 2 & 8: Ovine 8039, segment RVD1, day 1

Lane 3 & 9: Ovine 8038, segment LC, day 3

Lane 4 & 10: Ovine 8039, segment LC, day 3

Lane 5 & 11: Ovine 8038, segment RVD2, day 7

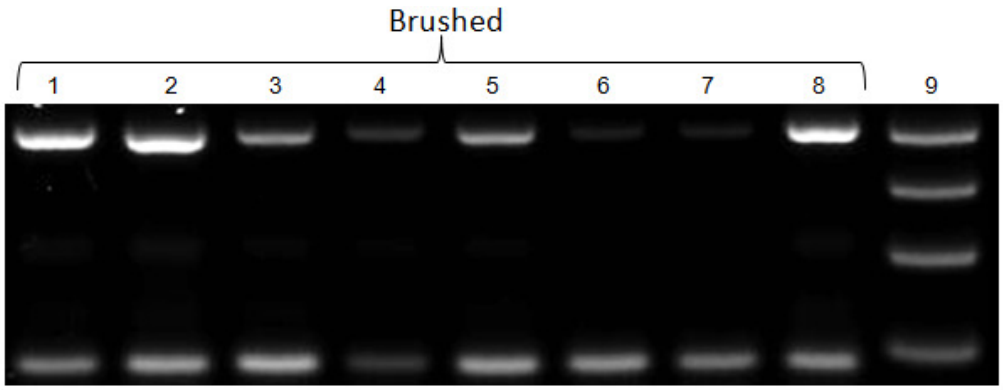
Lane 6 & 12: Ovine 8039, segment LVD1, day 7

Lane 13 & 14: Ovine 8038, segment RVD1, naive (duplicate)

Lane 15 & 16: Ovine 8039, segment RC, naive (duplicate)

Lane 17 : PCR negative control

Lane 18 : 1 kb DNA ladder



(*cont…*) Duplex RT-PCR products of KRT5 and ATPase genes amplified from brushed samples at days 7, 14, 21 and 28 post injuries

Lane 1: Ovine 2683, segment LVD2, day 7

Lane 2: Ovine 2683, segment RC, day 7

Lane 3: Ovine 2683, segment LVD1, day 14

Lane 4: Ovine 2683, segment RA, day 14

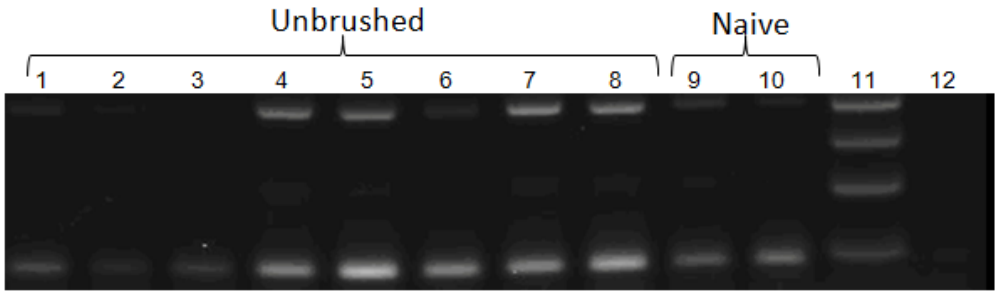
Lane 5: Ovine 2676, segment LVD2, day 21

Lane 6: Ovine 2676, segment RC, day 21

Lane 7: Ovine 2676, segment LVD1, day 28

Lane 8: Ovine 2676, segment RA, day 28

Lane 9: 1kb DNA ladder



(*cont…*) Duplex RT-PCR products of KRT5 and ATPase genes amplified from unbrushed samples at days 7, 14, 21 and 28 post injuries

Lane 1: Ovine 2683, segment LVD2, day 7

Lane 2: Ovine 2683, segment RC, day 7

Lane 3: Ovine 2683, segment LVD1, day 14

Lane 4: Ovine 2683, segment RA, day 14

Lane 5: Ovine 2676, segment LVD2, day 21

Lane 6: Ovine 2676, segment RC, day 21

Lane 7: Ovine 2676, segment LVD1, day 28

Lane 8: Ovine 2676, segment RA, day 28

Lane 9: Ovine 2683, segment RCD, naïve (d14)

Lane 10: Ovine 2683, segment RCD, naïve (d28)

Lane 11: 1kb DNA ladder

Lane 12: PCR negative control

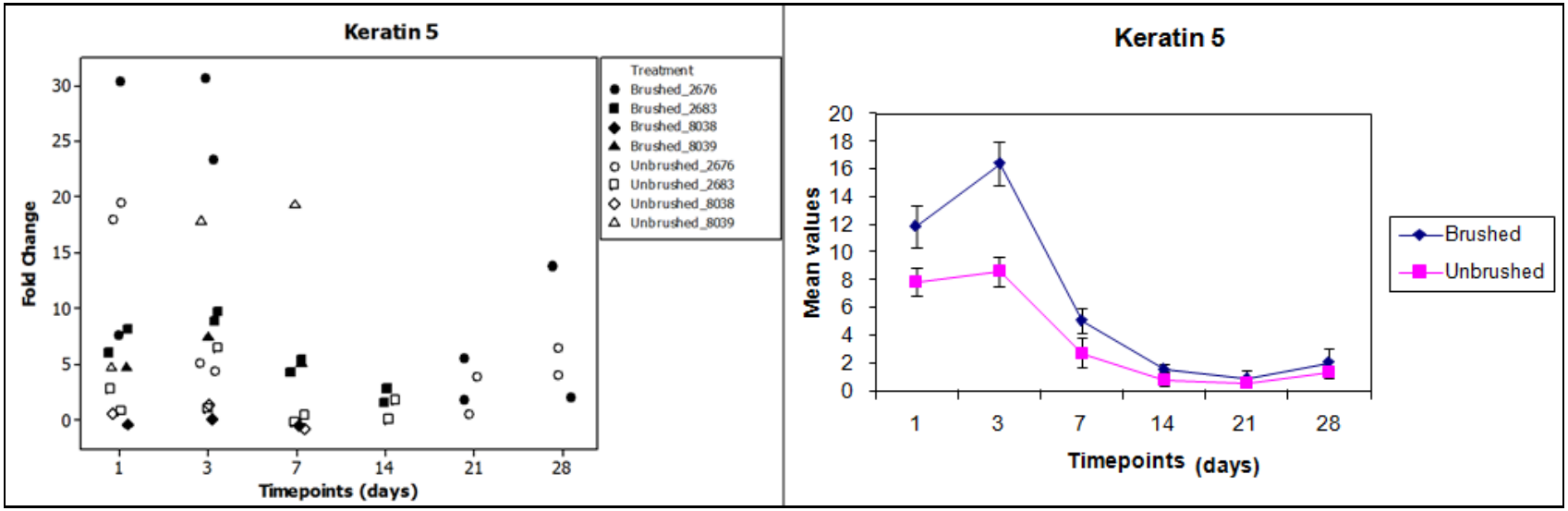


Figure 3.6(b): Graphs depicting the expression patterns of KRT5 between 1 and 28 days after injury. Graphs were plotted based on the fold change of the treated airways (brushed and unbrushed) relative to their respective naive airway (left panel) and based on the mean values (±SD). The expression profile observed indicates that KRT5 is up-regulated throughout the initial time points (right panel) whilst whilst the individual fold changes plots show a considerable variation in expression of this gene (left panel).

Changes in the expression of keratin 14 (where the underlining reflects the uncertainty regarding the true identity of the target in this instance) appeared to follow a similar pattern of initial (days 1 to 3) upregulation (relative to ATPase) followed by a decrease through day 7 to relatively consistent levels of expression by day 14 onwards (Figure 3.7(b); right panel). Although there was the suggestion of marginally increased levels of expression in brushed relative to unbrushed sites the individual value plots (Figure 3.7(b); left panel) indicated the considerable variability within and between sheep and cautioned against over-interpreting this data.

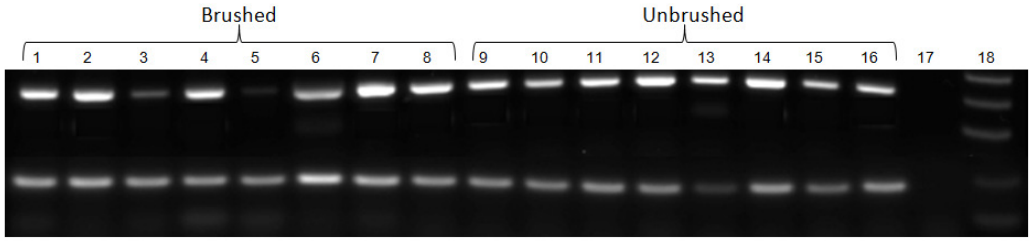


Figure 3.7(a): Agarose gel for a semi quantitative RT-PCR analysis for the keratin 14 gene
RT-PCR products of KRT14 and ATPase genes amplified from brushed (lanes 1-8) and unbrushed (lanes 9-16) samples at days 1 and 3 post injuries, and a PCR negative control (lane 17).

Lane 1 & 9: Ovine 2683, segment LVD1, day 1

Lane 2 & 10: Ovine 2683, segment LCD, day 1

Lane 3 & 11: Ovine 2676, segment LCD, day 1

Lane 4 & 12: Ovine 2676, segment RVD1, day 1

Lane 5 & 13: Ovine 2683, segment RVD2, day 3

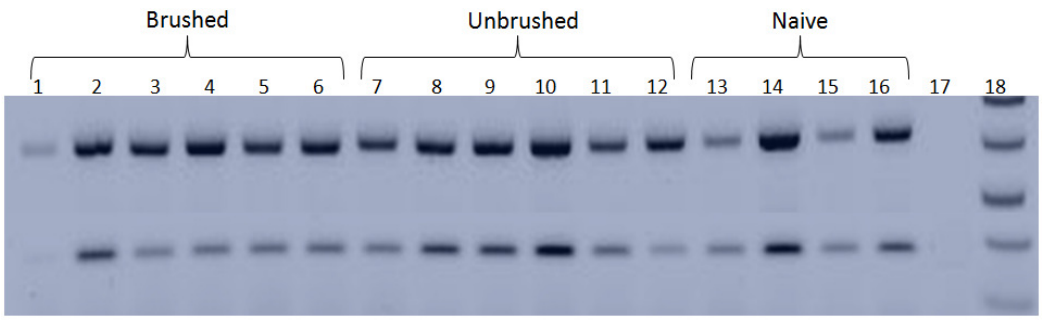
Lane 6 & 14: Ovine 2683, segment LC, day 3

Lane 7 & 15: Ovine 2676, segment LC, day 3

Lane 8 & 16: Ovine 2676, segment RVD2, day 3

Lane 17 : PCR negative control

Lane 18 : 1 kb DNA ladder



(...cont) **Agarose gel for a semi quantitative RT-PCR analysis for keratin 14**
Duplex RT-PCR products of KRT14 and ATPase genes amplified from brushed (lanes 1-6) and unbrushed (lanes 7-12) samples at days 1, 3 and 7 post injury, and a PCR negative control (lane 17).

Lane 1 & 7: Ovine 8038, segment RC, day 1

Lane 2 & 8: Ovine 8039, segment RVD1, day 1

Lane 3 & 9: Ovine 8038, segment LC, day 3

Lane 4 & 10: Ovine 8039, segment LC, day 3

Lane 5 & 11: Ovine 8038, segment RVD2, day 7

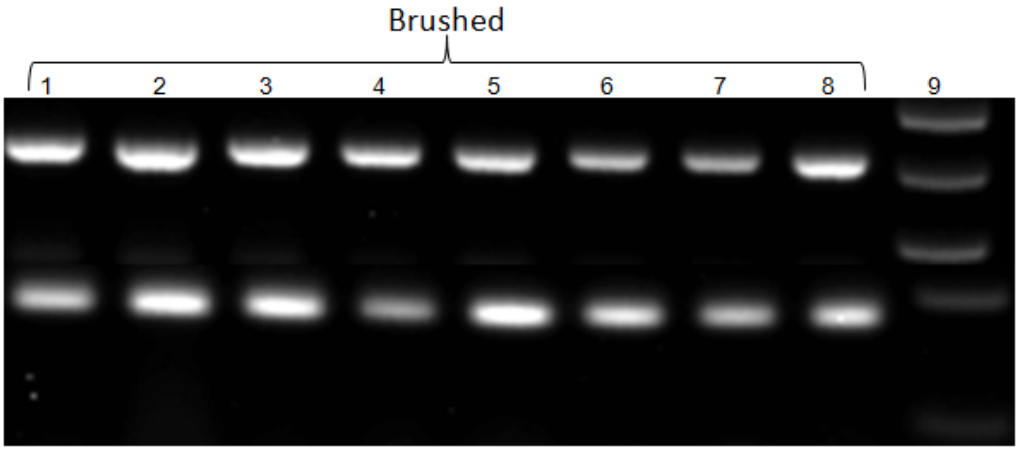
Lane 6 & 12: Ovine 8039, segment LVD1, day 7

Lane 13 & 14: Ovine 8038, segment RVD1, naive (duplicate)

Lane 15 & 16: Ovine 8039, segment RC, naive (duplicate)

Lane 17 : PCR negative control

Lane 18 : 1 kb DNA ladder



(cont…) **RT-PCR products of KRT14 and ATPase genes amplified from brushed samples at days 7, 14, 21 and 28 post injuries**

Lane 1: Ovine 2683, segment LVD2, day 7

Lane 2: Ovine 2683, segment RC, day 7

Lane 3: Ovine 2683, segment LVD1, day 14

Lane 4: Ovine 2683, segment RA, day 14

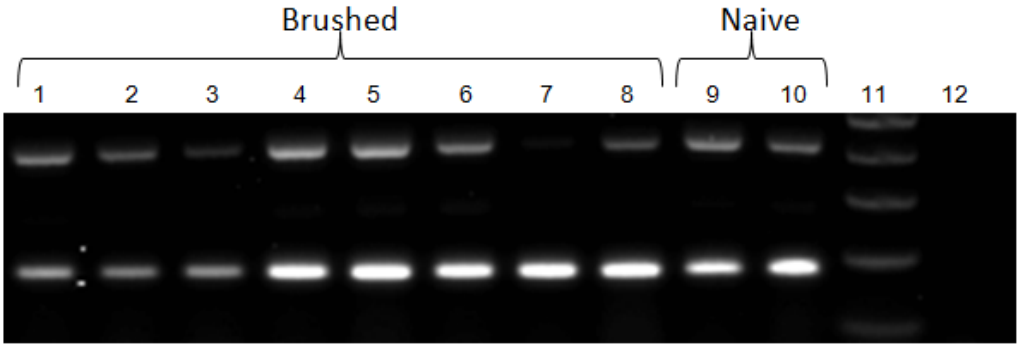
Lane 5: Ovine 2676, segment LVD2, day 21

Lane 6: Ovine 2676, segment RC, day 21

Lane 7: Ovine 2676, segment LVD1, day 28

Lane 8: Ovine 2676, segment RA, day 28

Lane 9: 1kb DNA ladder



(cont…) **RT-PCR products of KRT14 and ATPase genes amplified from unbrushed samples at days 7, 14, 21 and 28 post injuries**

Lane 1: Ovine 2683, segment LVD2, day 7

Lane 2: Ovine 2683, segment RC, day 7

Lane 3: Ovine 2683, segment LVD1, day 14

Lane 4: Ovine 2683, segment RA, day 14

Lane 5: Ovine 2676, segment LVD2, day 21

Lane 6: Ovine 2676, segment RC, day 21

Lane 7: Ovine 2676, segment LVD1, day 28

Lane 8: Ovine 2676, segment RA, day 28

Lane 9: Ovine 2683, segment RCD, naïve (d14)

Lane 10: Ovine 2683, segment RCD, naïve (d28)

Lane 11: 1kb DNA ladder

Lane 12: PCR negative control

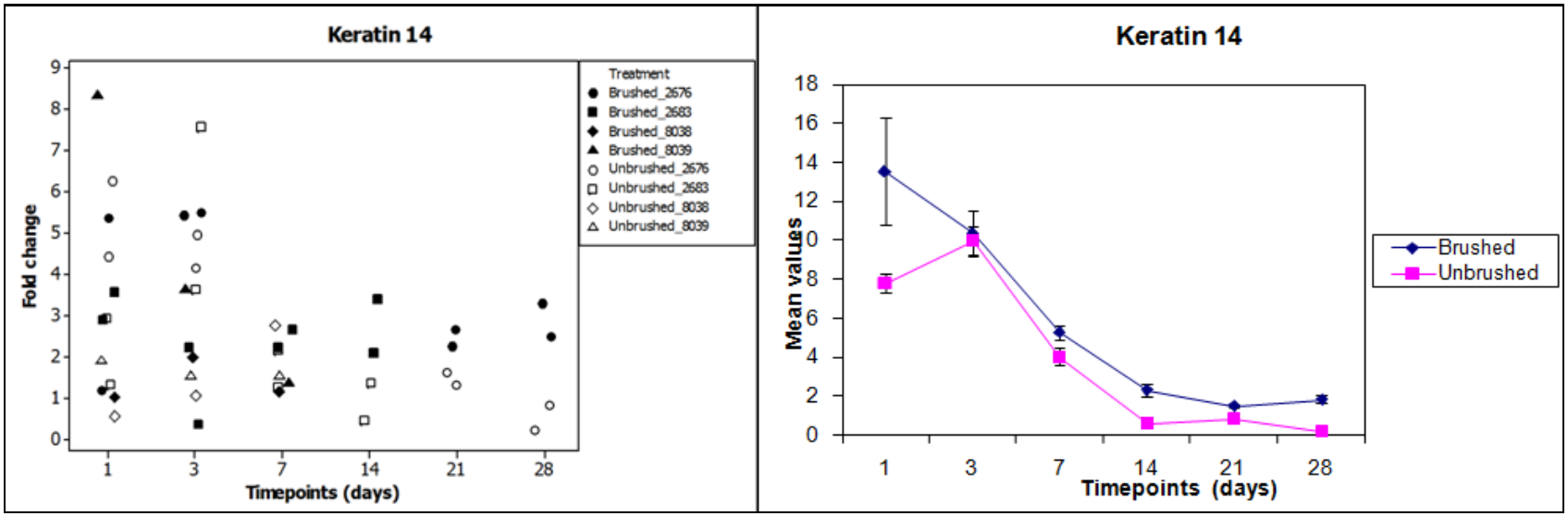


Figure 3.7(b): Graphs depicting the expression patterns of KRT14 between 1 and 28 days after injury. Graphs were plotted based on the fold change of the treated airways (brushed and unbrushed) relative to their respective naive airway (left panel) and based on the mean values (±SD). There is a marginally increased levels of expression in brushed relative to unbrushed sites and the individual value plots (right panel) indicated the considerable variability within and between sheep (left panel).

FoxJ1 gene expression appeared to vary according to the stage of repair, with the highest values being expressed initially followed by consistent levels at days 7 and 14 (Figure 3.8; right panel). Thereafter (days 21 and 28), expression in unbrushed sites exceeded that in brushed sites. As with the other candidate genes there was considerable variation in the fold-changes relative to naïve both within- and between-sheep (Figure 3.8(d); left panel).

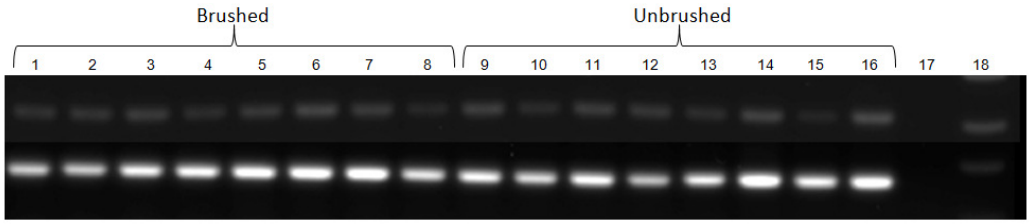


Figure 3.8(a): Agarose gel for a semi quantitative analysis of RT-PCR analysis for the FoxJ1 gene
RT-PCR products of the FoxJ1 and ATPase genes amplified from brushed (lanes 1-8) and unbrushed (lanes 9-16) samples at days 1 and 3 post injuries, and a PCR negative control (lane 17)

Lane 1 & 9: Ovine 2683, segment LVD1, day 1

Lane 2 & 10: Ovine 2683, segment LCD, day 1

Lane 3 & 11: Ovine 2676, segment LCD, day 1

Lane 4 & 12: Ovine 2676, segment RVD1, day 1

Lane 5 & 13: Ovine 2683, segment RVD2, day 3

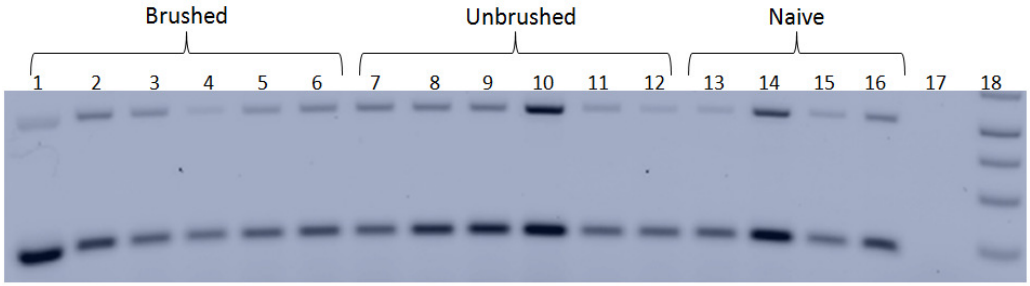
Lane 6 & 14: Ovine 2683, segment LC, day 3

Lane 7 & 15: Ovine 2676, segment LC, day 3

Lane 8 & 16: Ovine 2676, segment RVD2, day 3

Lane 17 : PCR negative control

Lane 18 : 1 kb DNA ladder



(...cont) **Agarose gel for a semi quantitative RT-PCR analysis for FoxJ1**
Duplex RT-PCR products of FoxJ1 and ATPase genes amplified from brushed (lanes 1-6) and unbrushed (lanes 7-12) samples at days 1, 3 and 7 post injury, and a PCR negative control (lane 17).

Lane 1 & 7: Ovine 8038, segment RC, day 1

Lane 2 & 8: Ovine 8039, segment RVD1, day 1

Lane 3 & 9: Ovine 8038, segment LC, day 3

Lane 4 & 10: Ovine 8039, segment LC, day 3

Lane 5 & 11: Ovine 8038, segment RVD2, day 7

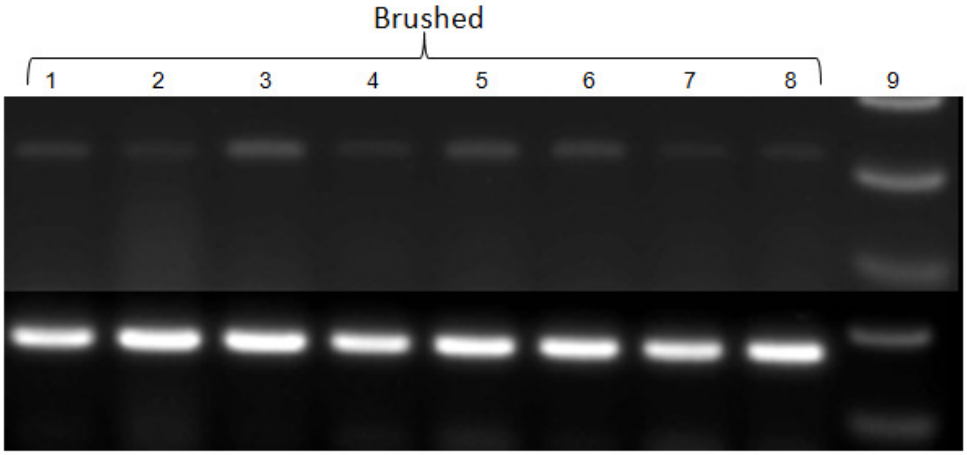
Lane 6 & 12: Ovine 8039, segment LVD1, day 7

Lane 13 & 14: Ovine 8038, segment RVD1, naive (duplicate)

Lane 15 & 16: Ovine 8039, segment RC, naive (duplicate)

Lane 17 : PCR negative control

Lane 18 : 1 kb DNA ladder



(cont…) **RT-PCR products of the FoxJ1 and ATPase genes amplified from brushed samples at days 7, 14, 21 and 28 post injuries**

Lane 1: Ovine 2683, segment LVD2, day 7

Lane 2: Ovine 2683, segment RC, day 7

Lane 3: Ovine 2683, segment LVD1, day 14

Lane 4: Ovine 2683, segment RA, day 14

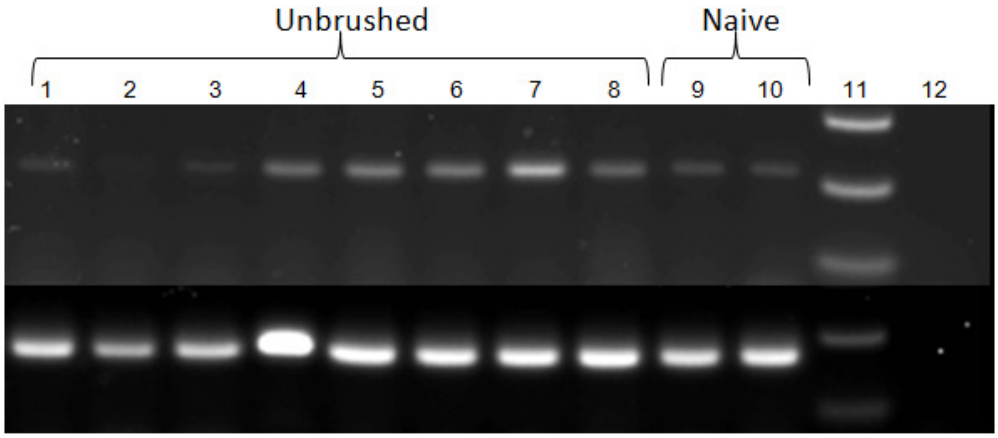
Lane 5: Ovine 2676, segment LVD2, day 21

Lane 6: Ovine 2676, segment RC, day 21

Lane 7: Ovine 2676, segment LVD1, day 28

Lane 8: Ovine 2676, segment RA, day 28

Lane 9: 1kb DNA ladder



(cont…) **RT-PCR products of the FoxJ1 and ATPase genes amplified from unbrushed samples at days 7, 14, 21 and 28 post injuries**

Lane 1: Ovine 2683, segment LVD2, day 7

Lane 2: Ovine 2683, segment RC, day 7

Lane 3: Ovine 2683, segment LVD1, day 14

Lane 4: Ovine 2683, segment RA, day 14

Lane 5: Ovine 2676, segment LVD2, day 21

Lane 6: Ovine 2676, segment RC, day 21

Lane 7: Ovine 2676, segment LVD1, day 28

Lane 8: Ovine 2676, segment RA, day 28

Lane 9: Ovine 2683, segment RCD, naïve (d14)

Lane 10: Ovine 2683, segment RCD, naïve (d28)

Lane 11: 1kb DNA ladder

Lane 12: PCR negative control

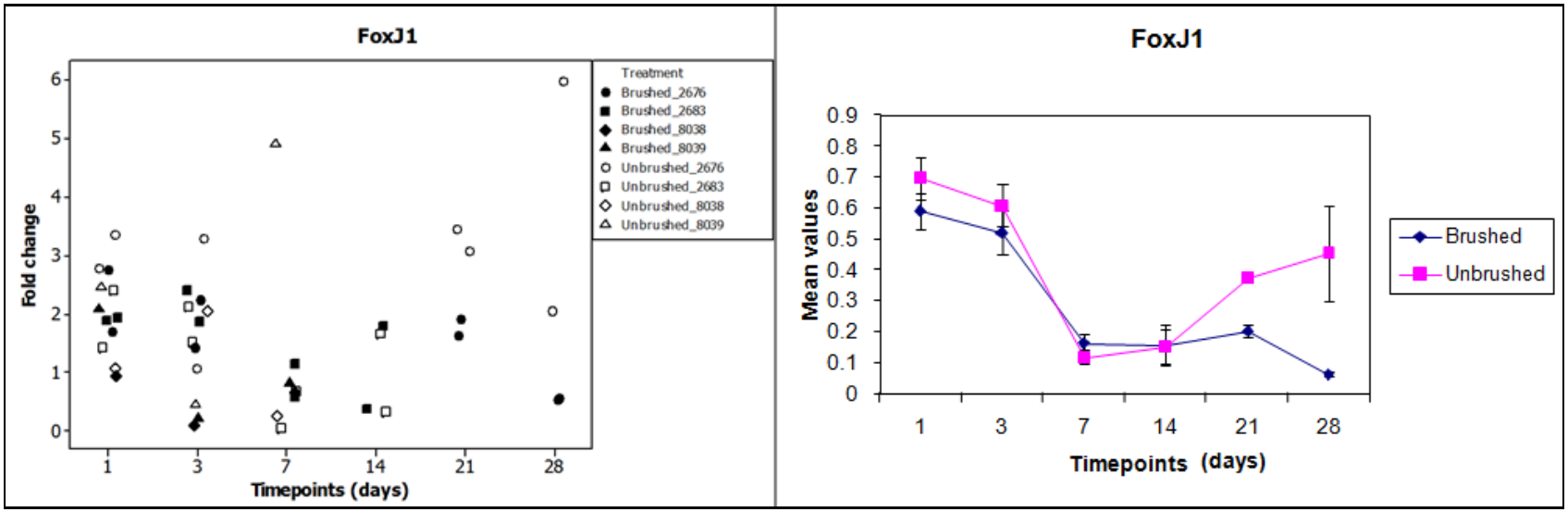


Figure 3.8(b): Graphs depicting the expression patterns of FoxJ1 between 1 and 28 days after injury. Graphs were plotted based on the fold change of the treated airways (brushed and unbrushed) relative to their respective naive airway (left panel) and based on the mean values (±SD). FoxJ1 gene expression appeared to vary according to the stage of repair especially on the earlier time points (right panel) whilst its a considerable variation in the fold-changes relative to naïve both within- and between-sheep (left panel).

For beta-tubulin, patterns of expression over time were similar to those already observed – namely elevated expression (relative to ATPase) initially (days 1 and 3), followed by a decrease in that expression through day 7 to broadly consistent levels of expression from days 14 to 28 (Figure 3.9(b); right panel). Levels of expression in brushed areas appeared to exceed those in unbrushed areas over the initial phase of repair, with this phenomenon being also reflected in the individual fold-change plots (Figure 3.9(b); left panel).

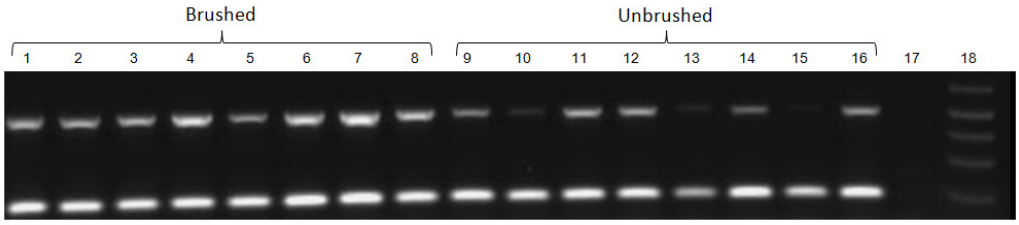


Figure 3.9(a): Agarose gel for a semi quantitative RT-PCR analysis for the beta-tubulin gene
Duplex RT-PCR products of beta-tubulin and ATPase genes amplified from brushed (lanes 1-8) and unbrushed (lanes 9-16) samples for days 1 and 3 post injuries, and a PCR negative control (lane 17).

Lane 1 & 9: Ovine 2683, segment LVD1, day 1

Lane 2 & 10: Ovine 2683, segment LCD, day 1

Lane 3 & 11: Ovine 2676, segment LCD, day 1

Lane 4 & 12: Ovine 2676, segment RVD1, day 1

Lane 5 & 13: Ovine 2683, segment RVD2, day 3

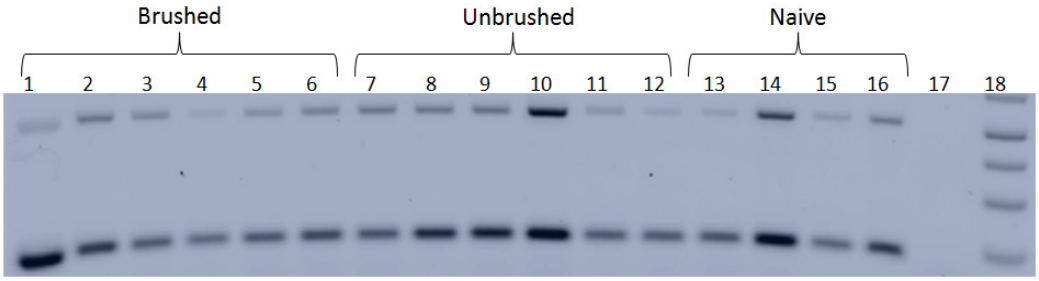
Lane 6 & 14: Ovine 2683, segment LC, day 3

Lane 7 & 15: Ovine 2676, segment LC, day 3

Lane 8 & 16: Ovine 2676, segment RVD2, day 3

Lane 17 : PCR negative control

Lane 18 : 1 kb DNA ladder



(...cont) **Agarose gel for a semi quantitative RT-PCR analysis for beta-tubulin**
Duplex RT-PCR products of TUBB and ATPase genes amplified from brushed (lanes 1-6) and unbrushed (lanes 7-12) samples at days 1, 3 and 7 post injury, and a PCR negative control (lane 17).

Lane 1 & 7: Ovine 8038, segment RC, day 1

Lane 2 & 8: Ovine 8039, segment RVD1, day 1

Lane 3 & 9: Ovine 8038, segment LC, day 3

Lane 4 & 10: Ovine 8039, segment LC, day 3

Lane 5 & 11: Ovine 8038, segment RVD2, day 7

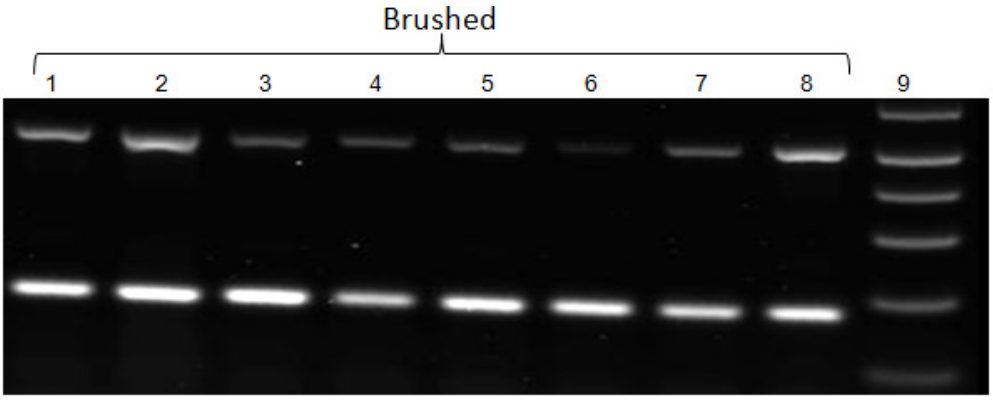
Lane 6 & 12: Ovine 8039, segment LVD1, day 7

Lane 13 & 14: Ovine 8038, segment RVD1, naive (duplicate)

Lane 15 & 16: Ovine 8039, segment RC, naive (duplicate)

Lane 17 : PCR negative control

Lane 18 : 1 kb DNA ladder



(cont…) **Duplex RT-PCR products of beta-tubulin and ATPase genes amplified from brushed samples for days 7, 14, 21 and 28 post injuries.**

Lane 1: Ovine 2683, segment LVD2, day 7

Lane 2: Ovine 2683, segment RC, day 7

Lane 3: Ovine 2683, segment LVD1, day 14

Lane 4: Ovine 2683, segment RA, day 14

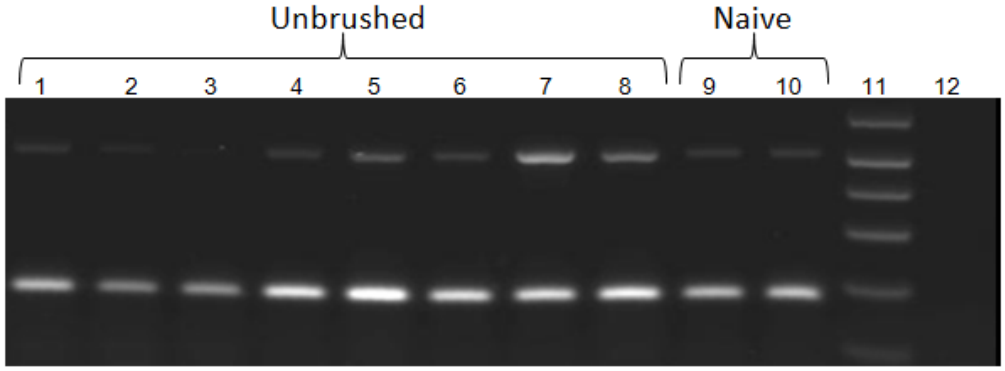
Lane 5: Ovine 2676, segment LVD2, day 21

Lane 6: Ovine 2676, segment RC, day 21

Lane 7: Ovine 2676, segment LVD1, day 28

Lane 8: Ovine 2676, segment RA, day 28

Lane 9: 1kb DNA ladder



(cont…) **Duplex RT-PCR products of beta-tubulin and ATPase genes amplified from unbrushed samples for days 7, 14, 21 and 28 post injuries, and a PCR negative control (lane 12).**

Lane 1: Ovine 2683, segment LVD2, day 7

Lane 2: Ovine 2683, segment RC, day 7

Lane 3: Ovine 2683, segment LVD1, day 14

Lane 4: Ovine 2683, segment RA, day 14

Lane 5: Ovine 2676, segment LVD2, day 21

Lane 6: Ovine 2676, segment RC, day 21

Lane 7: Ovine 2676, segment LVD1, day 28

Lane 8: Ovine 2676, segment RA, day 28

Lane 9: Ovine 2683, segment RCD, naïve (d14)

Lane 10: Ovine 2683, segment RCD, naïve (d28)

Lane 11: 1kb DNA ladder

Lane 12: PCR negative control

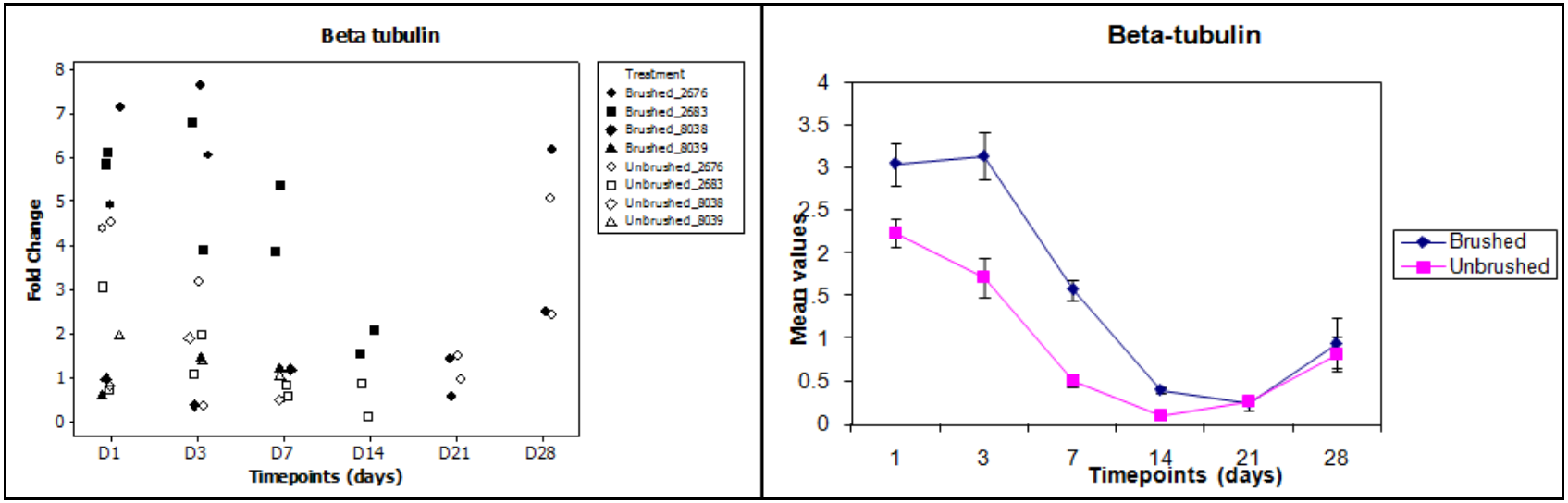


Figure 3.9(b): Graphs depicting the expression patterns of beta-tubulin between 1 and 28 days after injury. Graphs were plotted based on the fold change of the treated airways (brushed and unbrushed) relative to their respective naive airway (left panel) and based on the mean values (±SD). The expression level in brushed areas appeared to exceed those in unbrushed areas over the initial phase of repair (right panel) whilst and the individual value plots (right panel) indicated the considerable variability within and between sheep (left panel).

Considering the variation in expression patterns observed between samples, from different segments and/or sheep, but representing the same treatment condition, effort was directed towards establishing whether variable ATPase expression might contribute towards this phenomenon. The densitometry values for ATPase for each sample were plotted and (Figure 3.10) shows the expression profiles of ATPase gene for the beta-catenin and keratin 5 genes graphed according to the time points. Marked variation of ATPase expression levels were observed over the time course and these changes were not consistently reflected amongst samples deemed equivalent (Figure 3.10). Some of this variation may arise during PCR preparation. Indeed, with some of the genes being prepared using a singleplex PCR (in which ATPase and the genes of interest were prepared separately), variation might have occurred during preparation of the master mix i.e pipetting errors - especially in relation to the amount of template, primers and magnesium concentration.

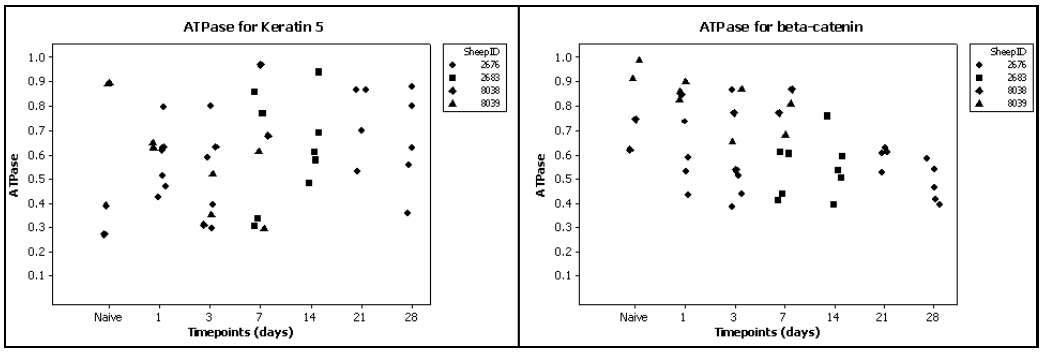


Figure 3.10: Scatter plots demonstrating the expression of ATPase throughout the time of injury for samples relating to keratin 5 (left panel) and beta-catenin (right panel).

3.5 DISCUSSION

As discussed in the previous chapter a large animal model has been developed that allows us to investigate the mechanisms of repair and regeneration of the airway following bronchial brushing. The studies presented in the present chapter were designed to build on the preliminary studies (Chapter 2) in order to assess the changes in expression of certain candidate genes, the expression of which are acknowledged to relate to cellular regeneration (basal and ciliated cells in particular) in the early and late processes of tissue repair. As discussed in the previous chapter, one limitation of the model is the lack of specific antibodies to identify and localise specific cell types that respond to airway injury.

This study adopted a molecular strategy to study the expression profiles associated with the regeneration and repair process following bronchial brushing. Although the ovine genome has not been completely sequenced, there is a substantial amount of ovine sequence information available in databases such as DFCI, EMBL-EBI and GenBank (although often as unclassified EST sequences). Therefore, the identification and classification of gene annotations rely on conventional bioinformatics analysis such as blasting through the databases with sequence from the other species (such as bovine, human and mouse). Thus, sequence alignment analysis was used to demonstrate the areas conserved between species. Resulting ovine sequences were checked for sequence similarity using GenBank.

The designed primers for the genes of interest used in this study were thus based on sequences conserved between species. Therefore the amplified PCR products were cloned and sequenced to confirm that the target sequence showed similarity to the sequence from which the primers were designed. The ovine sequences of these genes have now been made available via the network data servers of GenBank which are also available to EMBL in Europe and the DNA Data Bank of Japan for wider reference. It is not surprising to note that these gene (except keratin 14) sequences most closely resemble the bovine equivalents with similarity as high as 98%.
For keratin 14, the sequence analysis revealed that the PCR product demonstrated 95% similarity with the bovine EST sequence for keratin 14, 92% similarity with

human mRNA KRT14 sequence and 90% similarity with mouse KRT14 mRNA sequence. However, as the sequence demonstrated greater similarity with bovine keratin 17 the question arose as to whether these primers did indeed amplify the keratin 14 gene. Initial analysis to check the specificity of the primers showed that, both forward and reverse primer sequences were 100% similar to many bovine keratin genes such as KRT42, KRT17, KRT15 and KRT13. However, whilst both primers (forward and reverse) showed 100% similarity to bovine (NM_001166575.1) keratin 14 and the reverse primer was 100% similar to human (NM_000526.4) keratin 14 gene, the forward primer was only 95% similar to the human gene. Therefore, the possibility exists that this amplified sequence belongs to other types of keratin genes. Indeed analysis suggests that the designed primers would be expected to amplify PCR products of the same size for other types of keratin genes – including keratins 10, 13, 15, 17 and 42. However, alignment analysis (Appendix 3.1 and Appendix 3.2) shows that the sequence for amplified "KRT14" gene was highly similar to 13, 15, 17 and 42 (except keratin 10 – repeat sequence alignment was performed with keratin 10 excluded) in terms of the similarity of both forward and reverse primers. It is appreciated that further analysis is required in order to confirm the reliability of this sequence in representing the keratin 14 gene and until such time it is not possible to comment on the expression of keratin 14 with any confidence in this model.

Cytokeratin belongs to a group of protein known as intermediate filaments. Six groups of intermediate filaments exist based on their cellular origin and they form the cytoskeletal structure in all cells. These groups are acidic cytokeratins, basic cytokeratins, vimentin, neurofilaments (neurons), nuclear laminin (nuclear lamina) and nestin (stem cells). Cytokeratins are classified according to mass: low weight (acidic type 1 cytokeratins) and high weight (basic or neutral type II cytokeratins). In the context of possible alternative keratin genes amplified in our study, KRT10, KRT13 and KRT17 are considered low weight molecules which group with the KRT14 gene. In terms of epithelial cytokeratin expression: K10 is expressed in keratinized squamous epithelium, K13 is expressed in intermediate and superficial cells of non-keratinized mucosal stratified squamous epithelium, and KRT5, KRT14

and K17 are expressed in stratified epithelium. Therefore, although the identity of the amplified sequence in this study could not be accurately defined, the amplified product(s) still likely reflect genes that play a role in basal cell differentiation and may thus be relevant in the context of indicating potential involvement of basal cells in airway injury and repair.

In the present study, semi-quantitative RT-PCR assays were developed to investigate the expression profiles of the beta-catenin, beta-tubulin, keratins 5 and 14, and FoxJ1 genes in the airway wall following physical injury. Although limited numbers of sheep were used, qualitative appraisal suggested that the selected genes may be involved in regeneration and repair of the airway following bronchial brushing injury.

In our study using BrdU to label proliferative cells we observed increased proliferative activity at day 1 which reached a peak at day 3 when compared to the naïve and unbrushed airways (Chapter 2, Section 2.4.2.1). Histopathologically, at these time points there were some areas where the mucosa was still denuded at the brushed sites. From a molecular point of view, of the genes examined, only keratin 5 and beta tubulin showed a trend towards increased expression during the first 3 days post injury. Keratin 5 and keratin 14 are known as markers for basal cells (Boers JE et al., 1998, Voynow JA et al., 2005), and our finding of increased expression in brushed vs unbrushed sites suggests that these genes may play a role at a time when basal cells might be expected to contribute to the repair process. In cystic fibrosis airways cytokeratin 5/14 expression was found to be restricted to basal cells and to zones of hyperplasia (LeSimple P et al., 2007).

A rapid migration of basal cells at the area of damage is thought to be an important aspect of airway regeneration and repair. The histopathological and molecular aspects of this study support the hypothesis that basal-like cells might be involved in early migration and proliferation and may ultimately give rise to new differentiated cells. While basal cells may give rise to the other cell types, such as the ciliated cell (Hong KU et al., 2004b, Hong KU et al., 2004a, Voynow JA et al., 2005), there is

also speculation that cells residing close to the sub mucosal glands (Engelhardt JF, 2001), and cells from the blood (Abe S et al., 2004) may migrate and differentiate to become specialised epithelial cells.

In addition, it has been shown that basal cells often remain in the injured region of the wounded airway (Lane BP and Gordon R, 1974, Rennard SI et al., 1991), and release fibronectin which can serve as a chemoattractant for other airway epithelial cells to migrate to the area of injury (Rennard SI et al., 1991). In a study by Hong KU et al., (2004b) which investigated the hypothesis that basal cells represent a multipotent progenitor cell type for renewal of the injured tracheal epithelium of mice exposed to naphthalene, it was demonstrated that the injury resulted in a rapid induction of cytokeratin 14 expression among the majority of Griffonia simplicifolia I (GS1)-B4-reactive cells. Restoration of depleted secretory cells occurred after 6 days of recovery and was associated with regression of the basal cell hyperplasia, suggesting that these basal cells have the capacity for restoration of a fully differentiated epithelium (Hong KU et al., 2004b).

Patterns of beta-catenin gene expression in the first 3 days post injury suggested that there was a trend towards a high expression of beta catenin in both brushed and unbrushed areas. This finding was in agreement with the finding reported in mice that were subjected to bronchiolar injury with naphthalene (Park KS et al., 2006). Both nuclear and cytoplasmic staining for beta-catenin was markedly increased as early as 24 to 48 hours post injury in the squamous and cuboidal cells lining the bronchioles (in the normal adult lung, beta-catenin is expressed on the cell membrane and is rarely observed in the nuclei of the airway epithelial cells) (Park KS et al., 2006). Protein expression was decreased (below normal) at four days post injury, and thereafter restored to the pattern seen in the normal lung. This study demonstrated that beta-catenin gene was able to translocate to the nucleus, and suggested that it may be associated with the activation of the transcriptional genes that are required for the later stages of airway remodelling and repair, including cellular differentiation.

Previous studies have suggested that beta-catenin plays a major role during lung development (Mucenski ML et al., 2003, Mucenski ML et al., 2005, Li C et al., 2002, De Langhe SP and Reynolds SD, 2008), however it is unclear whether this gene product plays a role in maintaining and regulating adult lung stem and/or progenitor cells (Stripp BR and Reynolds SD, 2008). It is likely that many signalling pathways interact to mediate normal regulation of the proliferative and differentiation responses of a tissue stem cell and roles for these signals in regulation of the reparative characteristics of different organs are likely to be highly tissue specific (Stripp BR and Reynolds SD, 2008). However, Reynolds SD et al., (2008) suggested that beta-catenin signalling functions primarily as a negative regulator of bronchiolar cell differentiation and attenuation of beta-catenin signalling during the late prenatal/early postnatal period promotes establishment of the bronchiolar stem cell hierarchy by promoting transit-amplifying cell (TAC) differentiation (Reynolds SD et al., 2008).

The expression of beta-tubulin and FoxJ1 genes were reported to be associated with the regeneration of ciliated cells in lung development and repair (Park KS et al., 2006). This is the first study to generate an ovine mRNA sequence for the FoxJ1 gene. Preliminary findings presented here suggest that expression of this gene changed according to the stage of the repair process. An *in vitro* study by You Y et al., (2004) with a primary air-liquid interface (ALI) culture model of wild-type mouse airway epithelial cells, indicated that FoxJ1 expression was present at day 2 of culture (in relatively undifferentiated cells) before the appearance of cilia and expression of beta-tubulin 1V on the cell surface. In contrast, apical beta-tubulin IV expression was not observed in the FoxJ1 null cells maintained under identical culture conditions, indicating a requirement for FoxJ1 for ciliogenesis (You Y et al., 2004).

Although there were similar patterns of expression between brushed and unbrushed areas, the expression of beta-tubulin at brushed sites exceeded that seen in unbrushed sites from days 1 to 14 following bronchial brushing. The pathophysiological changes observed on the tissue sections at days 7 and 14 post injury (Chapter 2)

suggest that the injury site is mostly covered with regenerated epithelial cells. However some areas of this epithelial layer were still not fully differentiated and a lack of ciliated cells was observed. Since the repair process involves migration and differentiation of the regenerated epithelium towards the central area of the injury, some areas (transitional areas) had partially or even fully differentiated epithelium at the earlier time points (days 1 to 3), thus it is possible that the expression of beta-tubulin and FoxJ1 is associated with epithelial cell differentiation in giving rise to ciliated cells at these areas in the initial stages of repair. Therefore, it could be that the variation in terms of the expression profiles observed results from the lack of a coordinated change in expression across the entire injury site.

The results of this study neither permit elucidation of the role of beta-tubulin and FoxJ1 during regeneration of the ciliated cells nor indicate how they cooperate. A study by Blatt EN et al., (1999) in embryonic mouse lung development revealed that HFH-4 (FoxJ1) protein expression is restricted to ciliated cells and temporally related to ciliogenesis. Characterization of the expression of this gene in developing organs suggests that FoxJ1 plays a role in regulation of the genes important for epithelial cell differentiation and ciliogenesis (Blatt EN et al., 1999). Taken together, it is likely that activated FoxJ1 is important in providing a foundation for the beta-tubulin gene to be expressed by fully differentiated, ciliated cells in a repaired epithelium.

Although this preliminary study used a limited number of sheep, the expression profiles of those genes observed in this study provide hints on how those genes responded to the injury and repair. However, there are obvious limitations in using this type of semi-quantitative RT-PCR technique to quantify the expression of the candidate genes. The technique provides less accurate quantification when compared to the probe-based assays developed for real-time PCR techniques. In addition, the evaluation of changes at the molecular level does not allow the identification and location of cell types that express their protein products. The ability to identify the locations and types of cells expressing the specific protein-markers would be required for a clearer picture of the repair process.

Another limitation in this study is the low RNA concentrations isolated from the cartilaginous airway tissue samples. The quality measurement of the RNA used in this study relied on spectrophotometric absorbance (A_{260}/A_{280} ratios). We found that, although all of RNA samples used in this study generated acceptable A_{260}/A_{280} and A_{260}/A_{230} ratios, there was still variation observed in terms of the RNA concentrations. Although the initial amounts of RNA used in the cDNA synthesis and the amounts of cDNA used in the PCR amplification were kept as consistent as possible, the observed variation in the ATPase gene PCR products was still a concern since the expression level of the gene of interest was normalised according to the ATPase gene. The variation in the expression profiles of the studied genes observed in this study might be partly due to unexpected and unexplained variation in the expression of the housekeeping gene (Figure 3.10).

In this preliminary study, the expression patterns of the genes seem to suggest that there are time dependent changes in expression of some of these genes following injury to the ovine airway epithelium. Limitation with respect to experimental design (n=4) preclude definitive statements in this regard and highlight the need to extend the scope of these investigations towards a more involved protocol wherein the design permits robust statistical appraisal.

3.6 CONCLUSION OF THE CHAPTER

Semi-quantitative RT-PCR assays were developed to evaluate the expression profiles of the ovine beta-catenin, beta-tubulin, keratins 5 and 14, and FoxJ1 genes in airway epithelium following bronchial brushing to assess whether some of the genes reported to be involved in the repair process in other species may also be involved in the sheep. Despite the limited number of sheep used this study, the information obtained in terms of the expression profiles of the studied genes suggests that these genes may be involved in the repair of the airway.

The limitations of this study in terms of sheep numbers and the accuracy and reproducibility the PCR assays led us to the conclusion that candidate gene approaches in this context are naturally limited by the scope of selection, and as such benefit would accrue from taking advantage of a more comprehensive approach to transcriptional analysis in order to broaden our understanding of the fundamental mechanisms involved. The availability of an ovine array for microarray analysis [(ovine oligo array (15k), Agilent Technologies, UK)] gives the opportunity to screen the expression of 15000 probe targets in order to identify lists of the genes significantly regulated during the course of repair. The application of such technology requires high quality RNA samples for reproducibility and reliability of the data. As the RNA isolation technique is considered crucial to ensuring adequate RNA quality and integrity for such microarray analysis we considered it appropriate to give due regard to optimising this process in the context of samples derived from airways in the lung.

CHAPTER 4

AN ANALYSIS OF TIME DEPENDENT CHANGES ASSOCIATED WITH AIRWAY EPITHELIAL REPAIR IN SHEEP

4.1 INTRODUCTION

The airway epithelium is a pseudostratified layer consisting of specialised cell types that play a key role in defence and protecting the respiratory tract by combating inhaled toxicants and pathogens. The ability of the airway to respond to injury varies depending on the type of injury. In the case of asthma and COPD, repeat exposure of the airway epithelium to injury leads to remodelling of the airway tissue with potential impact on its role as a unique barrier against external damaging agent. Following extensive injury resulting in epithelial denudation the repair process not only involves regeneration of a new epithelial cell layer to cover the denuded area but also involves subsequent proliferation and differentiation to reconstitute a fully differentiated epithelium.

The process of cellular migration to cover the denuded area involves initial dedifferentiation of epithelial cells bordering the lesion. Such dedifferentiation reflects the need to revert to cell types with the capacity to migrate and form interactions or communication with other cells in order to progress airway repair (Evans MJ et al., 1990, Evans MJ et al., 1989, Evans MJ and Plopper CG, 1988). Such interaction and communication involves specific molecules which in turn controls the whole process of the repair. Effective and complete repair can thus be seen as a well coordinated and balanced series of events which likely show significant interdependence at several levels.

Amongst these events, cellular proliferation and re-differentiation are critical aspects and reflect the need for cells to ‘sense’ their environment such that their fate is appropriate in the context of repair. Such fate may involve differentiation to specialised cell types, cell repair, or cell death. The mechanisms involved during repair following injury include interaction between aspects of the inflammatory response, activation of innate immunity, migration, apoptosis, activation of signal transduction, differentiation, proliferation, metabolism and epithelial re-differentiation.

During the development of this thesis, preliminary studies had been carried out which involved studies directed towards understanding the pathophysiological and molecular aspects of airway repair (Chapters 2 and 3, respectively). Appreciating that conclusions derived from such initial experiments were necessarily limited as a consequence of the relatively few sheep involved and the pre-defined nature of the molecular analyses it was considered appropriate to direct attention towards designing a more robust experimental approach to realise the objective of characterising the cellular and molecular changes associated with airway regeneration and repair. With proliferation being a key aspect of repair and resolution following injury we sought to characterise this aspect of the response using Ki67 labelling. The advantage of using Ki67 as a proliferative cell marker is that it can label cells at all of the cell cycle phases (other than resting phase (G_0) and thus will potentially label all cells that are involved in proliferative activity. Additionally, with potential roles identified for mucus and/or secretory cells in contributing to the early steps of dedifferentiation and migration following injury (Keenan KP et al., 1982a, Keenan KP et al., 1982b, Keenan KP et al., 1982c, Lane BP and Gordon R, 1974, McDowell EM et al., 1979) and a stem cell niche described adjacent to the ducts of mucus-producing sub mucosal glands (Borthwick DW et al., 2001), we sought to determine whether there was any evidence for co-association between mucus-producing cells and Ki67-positive cells in the sheep airway following injury.

Therefore, in order to improve our understanding of the airway and cellular changes during epithelium repair as described in the Chapter 2, the study in this chapter was designed to give more comprehensive understanding of the pathophysiological changes of the airway during repair in response to bronchial brushing. More focus was directed towards the early processes of epithelial cell dedifferentiation, migration, proliferation and re-differentiation in order to confirm our preliminary findings as described in the Chapter 2. This would represent the first attempt to use bronchial brushing as a tool to induce airway epithelial perturbation in healthy sheep in order to understand the mechanisms involved in airway epithelial regeneration and repair.

4.2 OBJECTIVES

The aims of this chapter are:

1. To confirm histopathological changes associated with repair following bronchial brushing injury
2. To investigate the time course of proliferation and repair activities following lung injury using Ki67 as a proliferative marker.
3. To investigate the effect of bronchial brushing on the extent and distribution of airway wall tissue components and mucus-producing cells during the time course of the study and to relate any changes to concomitant cell proliferative activity using Ki67 as a marker.

4.3 MATERIALS AND METHODS

4.3.1 Experimental design

Eight sheep (n=8) [(breed = Suffolk Cross), (sex; Male = 4, Female = 4), (Age; means = 14.75 ± 0.45 months old), (bodyweight; means = 54.5 ± 3.74 kg)] were involved. Lung health was ascribed on the basis of bronchoscopic examination undertaken approximately 2 weeks (14.8 ± 6.2 days) prior to entry into the described experimental protocol. Such visual examination, in determining the absence of airway disease, was later confirmed by necropsy examination at gross and light microscopic levels. Bronchial brushings were performed on anaesthetised sheep as previously reported (Chapter 2; Sections 2.3.3, 2.3.4 (exclude BrdU administration) and 2.3.5) at predefined sites (approximately 15mm x 4mm located on the ventral aspect of each airway) within the major airways at staged time-points 6 hours, 1 day, 3 days and 7 days prior to euthanasia. Selection of these time points was based on our preliminary study (Chapter 2) which suggested that the major aspects of dedifferentiation, migration, proliferation and re-differentiation occur during this period.

Each sheep was subjected to a protocol which involved bronchial brushing on three occasions and two separate sites on each occasion were brushed at each time point so that changes at the histology level could be correlated with associated gene

expression profiles. Four different protocols were used and identified as 316, 716, 731 and 736 (Table 4.1) with two sheep assigned to each protocol. This number identifies the combination of time points for brushings in the respective sheep assigned to the protocol (i.e 316 corresponds to 3 days, 1 day and 6 hours). In this study, two naïve sites (from segments not subjected to any brushing procedure) were collected from each sheep for both histopathological analysis and subsequent RNA extraction for microarray profiling (n=8 of naive control). In addition to comparison with unbrushed naïve sites, the undamaged areas opposite the brushed sites were used for comparison.

Protocol	**731**			**316**			**736**			**716**			**731**			**316**			**736**			**716**		
Sheep	**8039**			**45**			**8064**			**7958**			**8038**			**8003**			**7954**			**8020**		
Segment	-7d	-3d	-1d	-3d	-1d	-6h	-7d	-3d	-6h	-7d	-1d	-6h	-7d	-3d	-1d	-3d	-1d	-6h	-7d	-3d	-6h	-7d	-1d	-6h
RC				BBr				BBr							BBr			BBr			BBr		BBr	
LC		BBr			BBr					BBr				BBr					BBr				**BBr**	
RVD1			BBr			BBr					BBr						BBr		**BBr**					BBr
LVD1	BBr							**BBr**				BBr		**BBr**		BBr					**BBr**			
RVD2			**BBr**		**BBr**		BBr			**BBr**			BBr									BBr		
LVD2		**BBr**				**BBr**			BBr						**BBr**	**BBr**						**BBr**		
RCD	**BBr**						**BBr**					**BBr**	**BBr**					**BBr**		BBr				
LCD				**BBr**					**BBr**		**BBr**						**BBr**			**BBr**				**BBr**

Table 4.1: Table indicating the combination of brushing protocols applied in the present study.
Each sheep was subjected to a protocol which involved BBr on three occasions with two separate sites being brushed on each occasion. Sheep were assigned to one of four different possible combinations of times used for BBr prior to necropsy. Such protocols were identified as 316, 716, 731 and 736 with two sheep assigned to each protocol. This number identifies the combination of time points for BBr in the respective sheep assigned to the protocol. Grey-coloured boxes on each sheep indicate which segments were used as naïve controls areas in each lung. (-) prior to post mortem, d = day, h = hour
* BBr = collected for histology and BBr = collected for microarray; grey-boxed with lines indicate the naive segment allocated for histology experiment

4.3.2 Tissue samples and processing

At post mortem, the lung segments were carefully mapped and dissected out to identify the injured sites, and the segments allocated for microarray and histopathology experiments were separated. The tissues used for histological analysis were fixed with 10% neutral buffered formalin for 12 hours before being dehydrated in alcohol and embedded in paraffin.

4.3.3 Analysis of pathophysiological changes of the airways and IHC

Individual airway sections from each location were subjected to the following: 1) Standard HE staining; 2) IHC for Ki67 followed the procedure described for BrdU incorporation (Chapter 2; Section 2.3.6.2). Mouse monoclonal anti human Ki-67 (clone MM1) (Novocastra, UK) was used at a dilution of 1:100. 3) Combined AB-PAS stain as previously described (Chapter 2; Section 2.3.7).

4.3.3.1 Morphometric analysis

4.3.3.1.1 HE and Ki67 staining

The protocol for morphometric analysis of the tissues stained with HE and anti-Ki67 antibody was similar to that described for the group of sheep used in the study discussed in the Chapter 2 (Sections 2.3.8 and 2.3.8.1). However, additional components were added to the morphometric analysis of Ki67 staining. First, unlike morphometric analysis in the Chapter 2, in this study, submucosal gland tissue was also subjected to morphometric analyses on HE-stained section and IHC for Ki67 staining. Second, to compare the distribution of Ki67-positive cells on each tissue category, the number of positive cells residing in each tissue category was counted in addition to counting the number of points landing on each tissue type. Therefore, the proportion of Ki67-positive cells residing in each tissue category was measured.

4.3.3.1.2 AB-PAS staining

The principles underlying the morphometric analysis of AB-PAS staining was similar to that described for BrdU and Ki67 staining. The areas of interest were defined as before - damaged, transitional and undamaged. Only two tissue categories are associated with mucus producing cells – the epithelium contains goblet cells and

the glands contain serous cells. The number of AB-PAS positive cells on these tissues was estimated by counting the number of grid points landing on these cell types for each tissue category. In the mucosal compartment counts were obtained for epithelium and gland tissue whereas in the submucosal compartment only gland tissue was involved.

4.3.3.1.3 Statistical analysis

In these analyses (HE, AB-PAS and Ki67 staining), the sheep's identity (SheepID) was used as a random effect in order to control for any pseudoreplication, and to minimise the noise associated with variation between observations and/or measurements on the same sheep (*see* Chapter 2; Section 2.3.8.2). In order to assess the distribution of Ki67-positive cells between compartments, areas of interest and time points, a model with Poisson error was used whilst to assess the proportion of positive cells on each tissue category, the proportion was estimated by dividing the number of positive cells on the tissue measured with the number of points falling on that particular tissue (binomial error). A similar approach of binomial data was used to estimate the AB-PAS stained cells. By using the number of grid points falling on the tissues measured on the morphometric analysis for Ki67 staining, one could estimate the effect of the brushing procedure on the proportion of epithelial, glandular, smooth muscle and vascular tissue.

The data was analysed using the freely available statistical software environment R (The R Foundation, Software version 2.9.0, Vienna, Austria). In all cases, a $p<0.05$ were considered statistically significant.

4.4 RESULTS

4.4.1 Analysis of the histopathology of the airway

4.4.2 Model validation

Generalised linear mixed model (family binomial) ANOVA was applied to quantitative morphometric data obtained from HE-stained sections of undamaged sites in order to determine that airways selected from different lung segments shared similar structural morphology (*see* Chapter 2; Section 2.4.1.1). The right cranial

(RC) segment was used as a reference level. The results of this analysis indicated that in this model, with sheep identity included as a random effect and segment RC as a reference level, airway tissues from different segments did not differ significantly in their structural morphology ($p>0.05$) – (Figure 4.1). The details of the R ouputs are shown in the Supplement 3 on CD. In order to correlate the original data points plotted for the graph in the Figure 4.1 in terms of how the segments been selected for BBr, the Table 4.2 shows the numbers of each segment subjected to BBr.

Segments	Numbers been selected for BBr*
RC	8
LC	7
RVD1	6
LVD1	5
RVD2	4
LVD2	1
RCD	1

Table 4.2: Table shows the segments and the numbers of times they have been selected for bronchial brushing. * The segments allocated for BBr from different sheep and time points

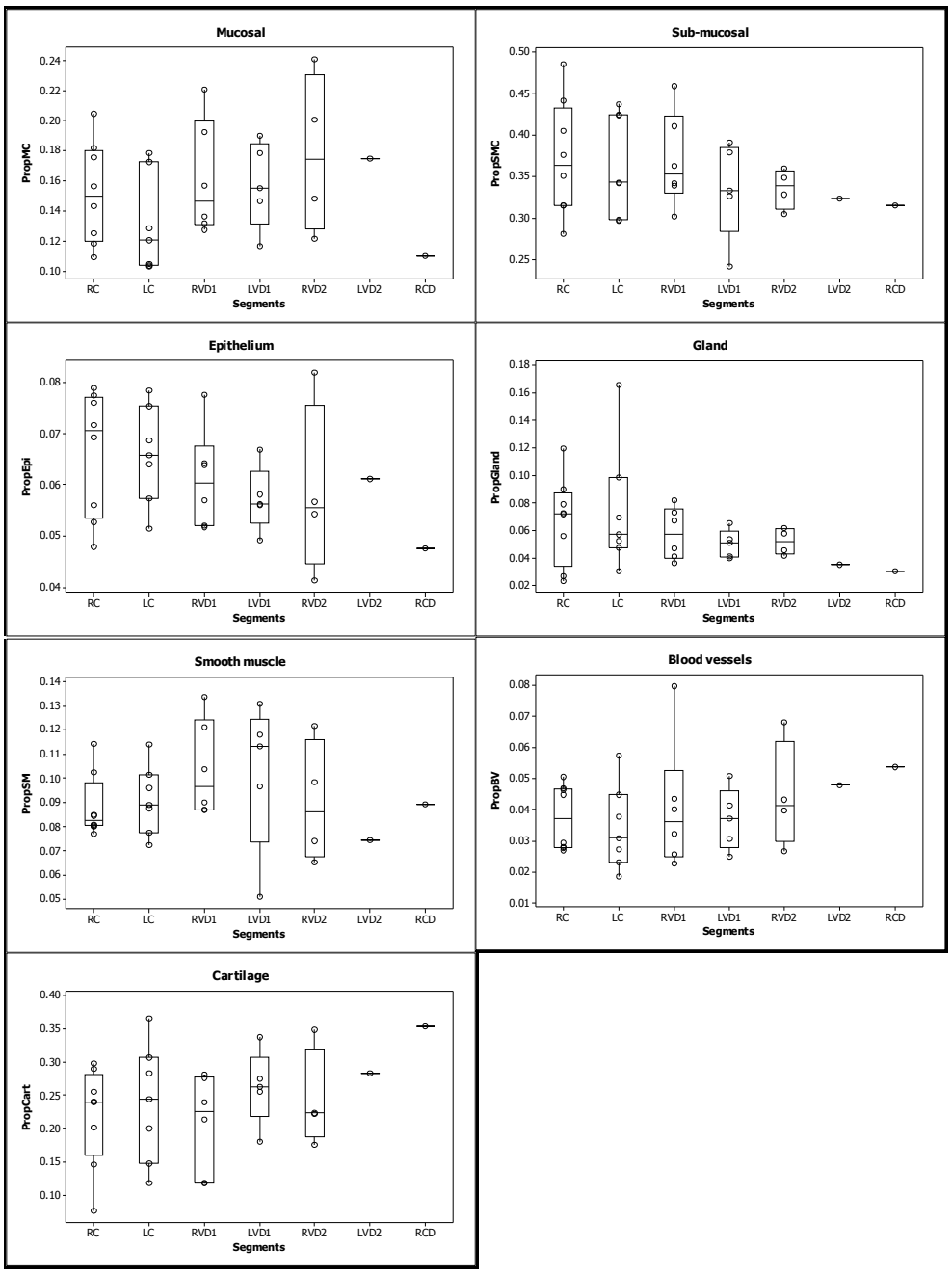


Figure 4.1: Box plots depicting the proportion of different tissue types found in the walls of undamaged airways derived from different lung segments
Plots illustrate the proportions of each tissue type lying within mucosal and submucosal compartments. There are no significant differences between segments (p>0.05). Symbols represent proportion data number of points landing on tissues for each segment measured. The numbers of points for each segment indicate the number of that particular segment subjected to histology experiment from different sheep i.e for RC segment, there are 8 airway sections of RC subjected to histology including the naive airway (Table 4.2) whilst only one sample from each of LVD2 and RCD was used for histology. Upper and lower box plot margins represent the interquartile range; middle bar indicates the median. The points outside the ends of the whiskers are outliers.

4.4.3 The effect of bronchial brushing on the relative proportion of airway tissues

Analysis by time point (areas of interest combined) showed that in brushed regions there was a significant decrease in the proportion of epithelium in the airway at early time points following injury (6 hours (mean proportion=0.103); p=0.017), day 1 (mean proportion= 0.110); p=0.029) but no significant difference at days 3 (p=0.939) and 7 (p=0.307) when compared to the naive (mean proportion=0.204) sites (mean proportion for both 0.2 and 0.2690, respectively) – (Figure 4.2; top left panel). If the analysis is done by area of interest (time points combined), then overall the damaged (mean proportion=0.053) sites have significantly less epithelium (p=0.000) than undamaged (mean proportion=0.224) but there was no significant difference between transitional (mean proportion=0.179) and undamaged (p=0.056).

Analysis by time point (areas of interest combined) demonstrated no significant effect of the brushing on the proportion of blood vessels at any of the time points (6 hours p=0.423 (mean proportion=0.089); day 1 p=0.671 (mean proportion=0.072); day 3 p=0.066 (mean proportion=0.105) and day 7 p=0.953 (mean proportion=0.077) as compared to time point 0 (mean proportion=0.078). If the analysis was completed according to area of interest (time points combined) the damaged area had significantly more blood vessels compared to the undamaged (p=0.0255) but not the transitional area (p=0.4737) – (Figure 4.2; top right panel) (mean proportion for damaged, transitional and undamaged were 0.097, 0.083 and 0.077, respectively). There was no effect of the brushing on the proportion of the gland structures in submucosal compartments (6 hour p =0.647 (mean proportion=0.089); day 1 p=0.134 (mean proportion=0.155); day 3 p=0.164 (mean proportion=0.151) and day 7 p=0.548 (mean proportion=0.122)) in all time points as compared to time point 0 (mean proportion=0.102) and in all three areas of interest (damaged p=0.837 (means proportion=0.113) and transitional p=0.225 (mean proportion=0.142)) – (Figure 4.2; bottom left panel) as compared to undamaged (mean proportion=0.117). The brushing had a significant effect on the proportion of smooth muscle at all time points in which less smooth muscle tissues were found at 6 hours (p= 0.0060; mean

proportion=0.202), day 1 (p=0.0123; mean proportion=0.214) and day 3 (p= 0.0199; mean proportion=0.222) following injury other than day 7 (p= 0.9306; mean proportion=0.346) as compared to the time point 0 (mean proportion=0.352). There was a significant reduction in proportion of the smooth muscle tissues in damaged (p=0.0026; mean proportion=0.203) and transitional (p=0.02; mean proportion=0.234) areas when compared to the undamaged area (mean proportion=0.312) – (Figure 4.2; bottom right panel). The details of the R outputs are shown in the Supplement 4 on CD.

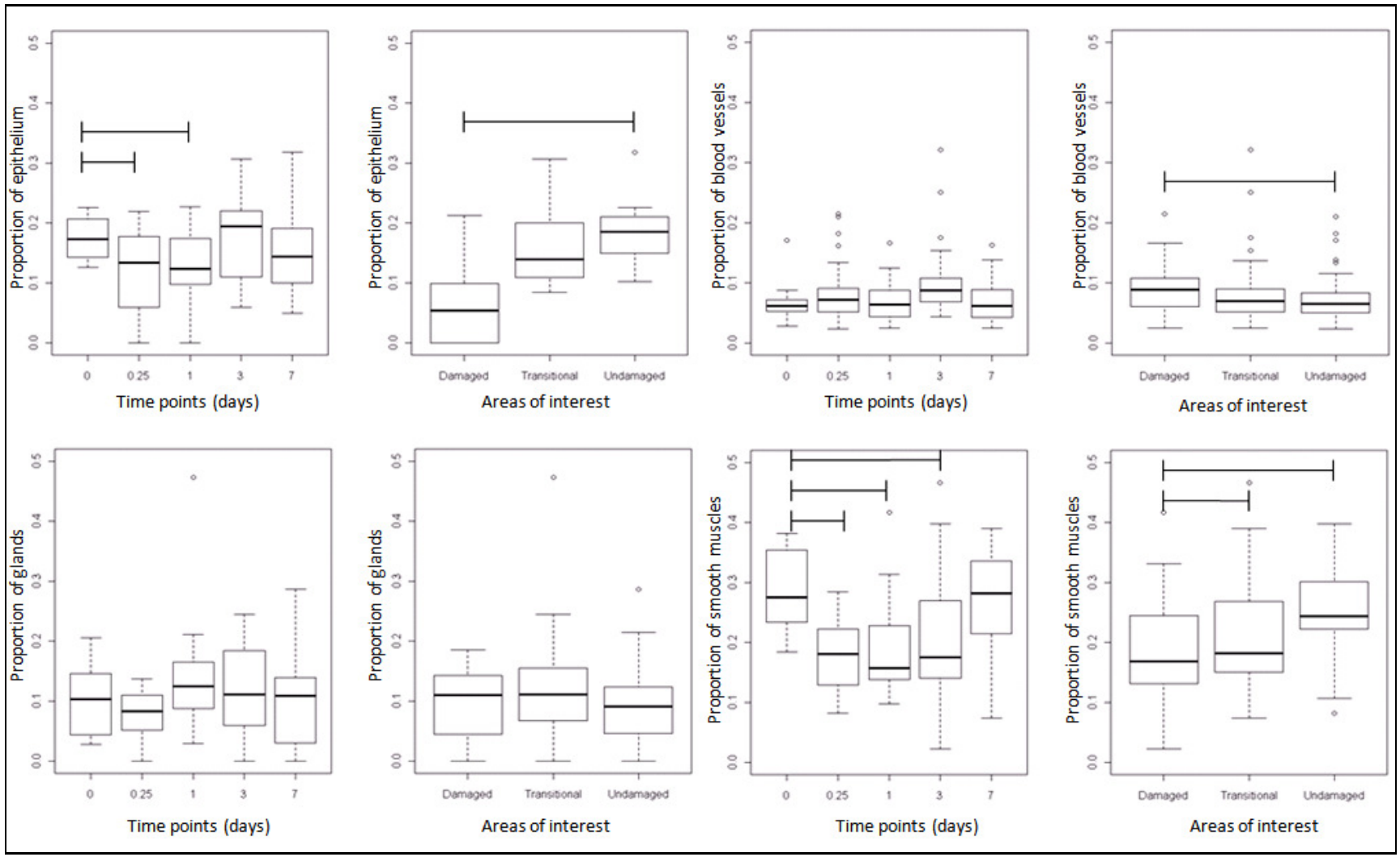

Proportion of epithelium
Time points (days)
Proportion of epithelium
Areas of interest
Proportion of blood vessels
Time points (days)
Proportion of blood vessels
Areas of interest
Proportion of glands
Time points (days)
Proportion of glands
Areas of interest
Proportion of smooth muscles
Time points (days)
Proportion of smooth muscles
Areas of interest
0
0.25
1
3
7
Damaged
Transitional
Undamaged

Figure 4.2: The effect of bronchial brushing on the airway tissues
Boxplots illustrate the proportions of the individual tissue category measured either by time points or by areas of interest. Y-axis shows proportion of tissues, and X-axis shows the time points or areas of interest. Symbols represent proportion data number of points landing on tissues for each area measured. Upper and lower box plot margins represent the interquartile range; middle bar indicates the median. The points outside the ends of the whiskers are outliers. There is a significant loss ($p<0.05$) of epithelium structure affected by the brushing in 6 hours and day 1 following injury as compared to the time point 0 and there is a significant different ($p<0.05$) in the proportion of the epithelial tissue in damaged area compared to the undamaged (top and left panel). There is no significant effect on the proportion of the blood vessels in any times of injury as compared to the time point 0 ($p>0.05$) and the proportion of the blood vessels is higher in the damaged area as compared to the undamaged ($p<0.05$) – top right panel. Brushing showed no significant effect on the submucosal gland structure at any time points following injury as well as to the areas of interest ($p>0.05$) – bottom left panel. There is less smooth muscle tissues at 6 hours, days 1 and 3 post injury as compared to the time point 0 as affected by the brushing as well as in the damaged and transitional areas as compared to the undamaged ($p<0.05$) – bottom right panel.

4.4.4 Morphology of the airway following a bronchial brushing

One control site from a naïve segment and injured sites from three other segments per sheep (n=8 sheep) had been dissected out for analysis similar to that described in Chapter 2. The brushing procedure disrupted the morphology of the mucosal area including the basement membrane as previously described in the Chapter 2. Generally, the histopathological changes of the airways from 6 hours to day 7 following bronchial brushing were similar to those described in Chapter 2 (Section 2.4.1.2). Similarly, the injury extended to smooth muscle tissue in the submucosal compartment. However, the submucosal gland tissue was not affected by the brushing (*see* Section 4.4.3 in this chapter). At the damaged area at 6 hours post-injury (Chapter 2; Figure 2.7), there was generally no epithelium remaining with the exception of a single layer of cells in places which may be the first signs of regeneration of the epithelial layer. In HE-stained sections from six hours post injury, the denuded area adjacent to the remaining epithelium was partially covered with squamous epithelial cells. This single layer of cells bordering the lesion is likely to be migrating de-differentiated cells derived from the epithelium bordering the lesion which are involved in the ongoing process of repair. This dedifferentiation process at the marginal areas of the damaged site was clearly seen on sections stained with the AB-PAS staining at this time point (*see* Section 4.4.6 of this chapter). The number and morphology of the goblet cells were changed dramatically following brushing as compared to the undamaged epithelium and in the very few remaining goblet cells, PAS staining was observed at the apex of the cells (*see* Section 4.4.6 of this chapter).

At day 1 post injury (Chapter 2; Figure 2.8), the flattened cells were extending towards the centre of the damaged area. Some of the peripheral damaged areas were covered with new epithelium at the transitional area and some ciliated cells were occasionally observed. However, the basement membrane was still absent in central areas of the lesion. The lesion area became smaller compared to that seen at 6 hours post injury and the size continued to reduce over the time of injury. Whilst AB-PAS staining was evident at the transitional area, there was more evidence of dedifferentiation of goblet cells/ciliated cells at the transition areas compared to 6

hours post injury (*see* Section 4.4.6 of this chapter). At day 3 post injury (Chapter 2; Figure 2.9), the entire mucosal injury site was covered with a stratified squamous epithelial layer, and in some areas there was still evidence of mucosal haemorrhage. This undifferentiated epithelial cell layer was thickened in places suggesting an epithelial hyperplasia. Few ciliated cells could be seen and although in some areas the numbers of AB-PAS staining cells had begun to increase, the overall number was still reduced compared to normal (*see* Section 4.4.6 of this chapter). The damaged area was also overlaid with what appeared to be a mucin-rich layer with a dense and compact accumulation of neutrophils held within and below the mucin - a phenomenon previously described in Chapter 2.

At seven days post injury (Chapter 2; Figure 2.10) the epithelium at the brushing site appeared intact however the submucosa was thickened with granulation tissue and fibrosis. At this time point, the morphology of the epithelium had still not fully differentiated. The number of goblet cells in the repairing epithelium had increased compared to the day 3 but were still not back to normal levels (*see* Section 4.4.6 of this chapter).

4.4.5 Ki67 to label proliferative cells

In this study mouse anti-human Ki67 antibody was used to label proliferating cells in response to injury. In the airway sections Ki67-labelled cells were identified by IHC at the site of injury at all time points. Only a few Ki67-positive cells were observed in the undamaged area and these tended to reside at the basal area of the epithelium, blood vessels and submucosal glands (Figure 4.3 (a)). Sparsely scattered Ki67-labelled cells were identified as early as six hours post-injury; in both the mucosa and submucosa, not only in the damaged areas but also in transitional and undamaged areas (Figure 4.3 (b)). There was however no significant difference (Mucosa p=0.8953 and Submucosa p=0.4160) at 6 hours compared to time point 0 (naïve). Thereafter, the specific location and number of proliferating cells varied with the time from injury [(Figure 4.3 (a)) to (Figure 4.3 (e))].

In general, the distribution patterns of the Ki67-positive cells were similar to that seen in the BrdU-stained airway sections (Chapter 2) especially on the airway tissues such as epithelium, glands and blood vessels. However, there were more proliferative cells observed when sections were stained with anti-Ki67 compared to the BrdU - even in the airway sections at the later time point (day 7). At day 1 (Figure 4.3 (c)), Ki67-labelled cells were visible within the airway epithelium in clusters on both sides of the injury. The distribution of Ki67-positive cells extended to the area underneath the damaged epithelium in the mucosal and submucosal compartments, and especially in the regions of the submucosal glands (SMGs). At day 3 (Figure 4.3 (d)), more Ki67-labelled cells were observed in the epithelium and SMGs compared to day 1 of injury. Although the extent of proliferation was higher compared to the undamaged, the number of positive cells at day 7 was reduced relative to day 3 post injury (Figure 4.3 (e)). Overall, with the exception of the 6 hours time point (p=0.455; means=7.193), in all time points (with areas of interest and compartments combined), there were significantly increased numbers of Ki67-positive cells present (day 1; p=0.000 (means=42.573), day 3; p=0.000 (means=42.137) and day 7; p=0.000 (means=48.550)) as compared to the time point 0 (means=4.635) (Figure 4.4).

The Table 4.3 shows the details of p-values generated for the distribution of Ki67-positive cells on each areas of interest over the time of injury. When the distribution of Ki67-positive cells was compared between compartments (data for the areas of interest and the time points combined), the submucosal (geometric means=55.962) compartment had significantly less (p=0.0322) Ki67-positive cells as compared to the mucosal (geometric means=52.831) compartment (data not shown). To assess the distribution of Ki67-positive cells between the areas of interest (with the time points and compartments combined), significantly more Ki67-positive cells were found in the damaged (p=0.000; geometric means=94.148) and transitional (p=0.000; geometric means=81.014) areas as compared to the undamaged (geometric means=19.470) area (data not shown). The distributions of Ki67-positive cells were then compared between areas of interest and the time points for each compartment (Figure 4.5). As compared to the naïve airway (time point 0), there were significant

differences (p<0.05) in the Ki67-positive cell distribution at days 1, 3 and 7 (except day 1 in mucosal compartment) post injury (Table 4.3). In the transitional area, the distributions of Ki67-positive cells differed significantly (p<0.05) at days 1, 3 and 7 post injury as compared to the naïve airway (Table 4.3). For the undamaged area there were significantly more Ki67-positive cells at days 1, 3 and 7 (p<0.05) in the mucosa whilst only at day 1 and 3 in the submucosal compartment (Table 4.3). In order to get better overview of the statistical analysis outputs generated for these morphometric analyses, the details of the geometric mean values are shown in the appendix.

As described in the morphometric analysis section of the Methods, the numbers of Ki67-positive cells within different tissue categories were also counted involving the epithelium, glands, blood vessels and smooth muscle. For this analysis, a statistical model was created by using the number of Ki67-positive cells as the response variable, the time points and areas of interest (damaged, transitional and undamaged) used as factors and the total number of grids on each tissue category used as a covariate – by taking into account the effect of the size of the tissues on the number of Ki67-positive cells observed. For blood vessels, the data for both mucosal and submucosal compartments were combined whilst for the gland tissues only the glands in the submucosal compartment were involved in the analysis due to the fact that only a few glands were observed within the mucosal compartment. The analysis for the epithelium and smooth muscle was performed on the mucosal and submucosal compartments, respectively. In all tissue studied, there were significantly increased in the numbers of Ki67-positive cells within the tissue category for both damaged and transitional areas (p<0.05) as compared to the undamaged areas (with the time points combined). The details of p-values generated from the statistical analysis for the distribution of Ki67-positive cells over the time points on the tissues of interest are shown in the Table 4.4 and the details of their geometric means are shown in the appendix. The analyses were investigated further for each area of interest and the results are as follows:

In the epithelium (Figure 4.6(A)), the number of Ki67-positive cells were increased significantly at days 1, 3 and 7 (p<0.05) compared to naïve, in both damaged and transitional areas. There was a small increase in the number of proliferating cells observed at day 1 in the undamaged area (p<0.05).

Similarly, the number of Ki67-positive cells on gland tissues was increased (p<0.05) at days 1, 3 and 7 in the damaged area compared to the time point 0 (naïve) whilst in the transitional area, the number was significantly increased (p<0.05) only at days 1 and 3. There was no significant different observed in the undamaged area at all time points (p>0.05) – (Figure 4.6(B)).

In the blood vessels (Figure 4.6(C)), Ki67-positive cells were significantly increased (p<0.05) at days 1, 3 and 7 post injury compared to naïve not only in damaged and transitional areas but also in undamaged areas. At 6 hours, there was a significant increase in Ki67-positive cells in the transitional areas compared to time point 0 (naive).

In the smooth muscle (Figure 4.6(D)), the number of Ki67 positive cells was increased in the damaged area at days 3 and 7 post injury (p<0.05) but only at day 7 (p<0.05) in the transitional area. There was no significant different in terms of the number of Ki67-positive cells between the time points in the undamaged area (p>0.05).

The details of the R outputs are shown in the Supplement 5 on CD.

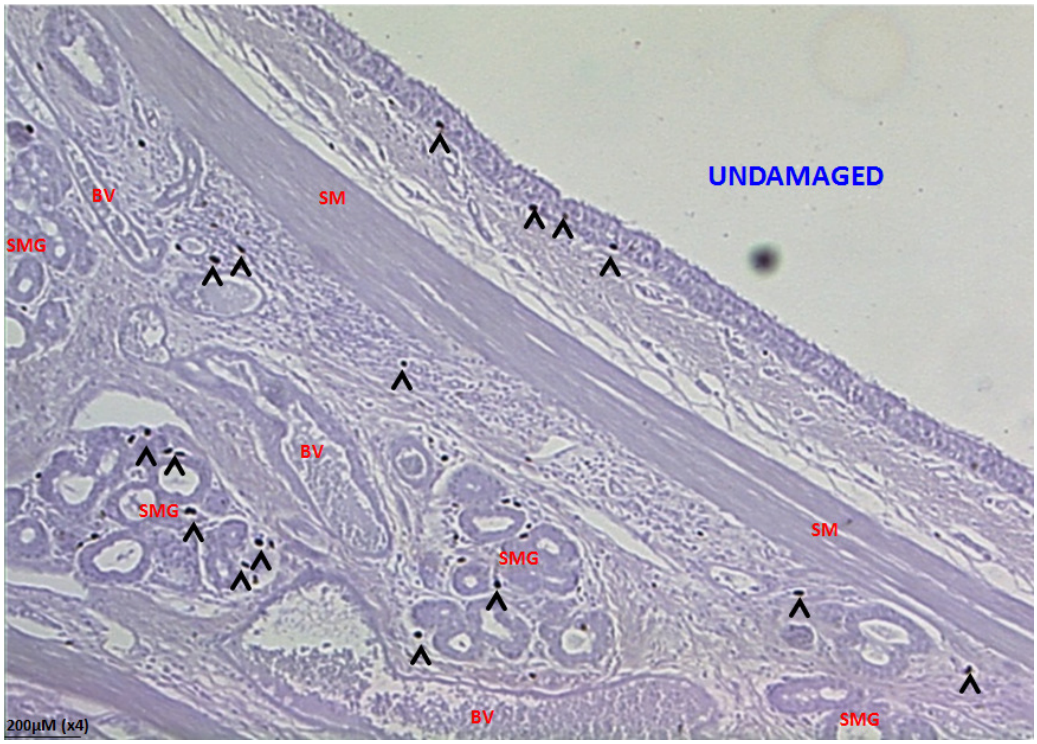


Figure 4.3(a): Undamaged airway tissue section stained with Ki67 antibody
A few numbers of Ki67-positive cells were observed in undamaged airway tissue sections. These cells were sparsely scattered along the basal area of the epithelium, and in association with glands and blood vessels in both mucosal and submucosal compartments. The arrows head show the Ki67-positive cells.

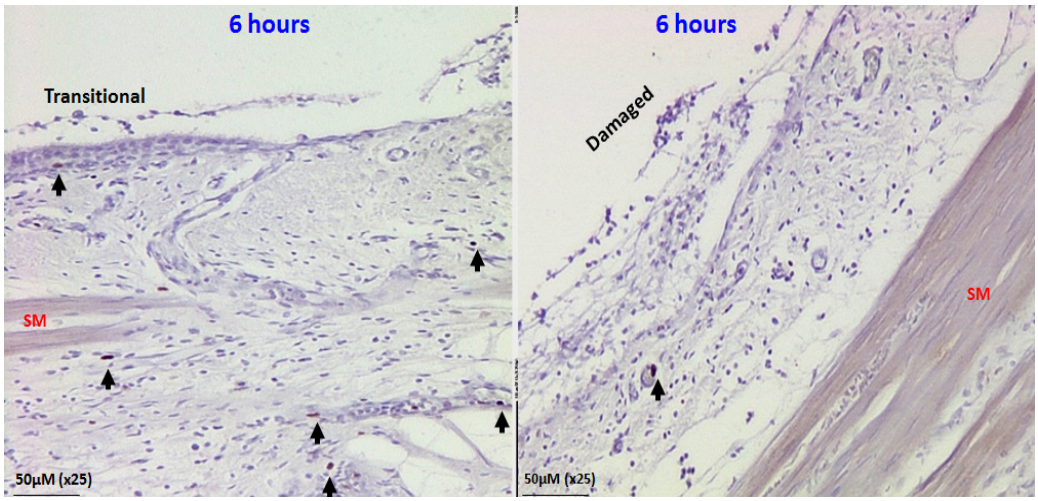


Figure 4.3 (b): Airway tissue section at 6 hours post injury stained with Ki67 antibody
At this time point, the positive stained cells were sparsely scattered within the transitional area and also underneath the damaged area. The distribution of these cells at this time point was not significantly different (p>0.05) to that found in the undamaged area. The arrows head show the Ki67-positive cells.

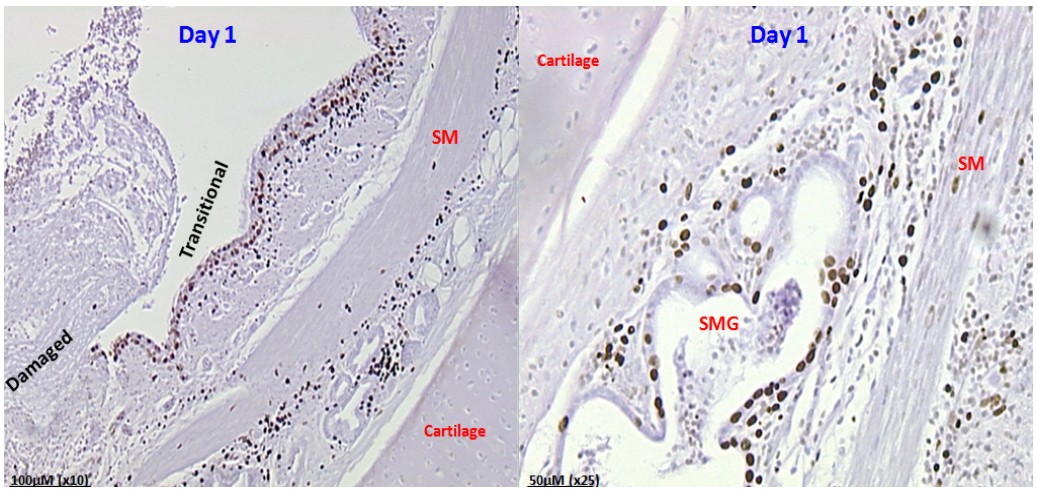


Figure 4.3(c): Airway tissue section at day 1 post injury stained with Ki67 antibody
Whilst the damaged site showed a clear mucosal disruption, the proliferative activity of the cells at this time point was increased significantly in response to injury. These cells accumulated at both edges of the lesion as well as in the submucosal glands. Different intensities of Ki67 staining reflect different stages of the cell cycle phase.

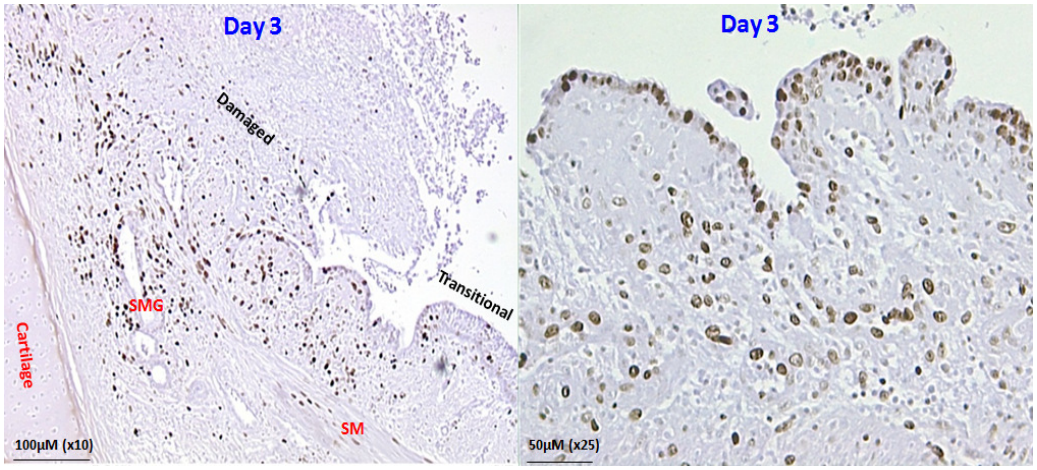


Figure 4.3(d): Airway tissue section at day 3 post injury stained with Ki67 antibody
The number of Ki67-positive cells observed at this time point was increased relative to day 1 with similar features observed at both edges of the lesion. In some sections, the damaged area was covered with new epithelium and stained cells accumulated at this area. Proliferative activity is still active in the submucosal gland region at this time point.

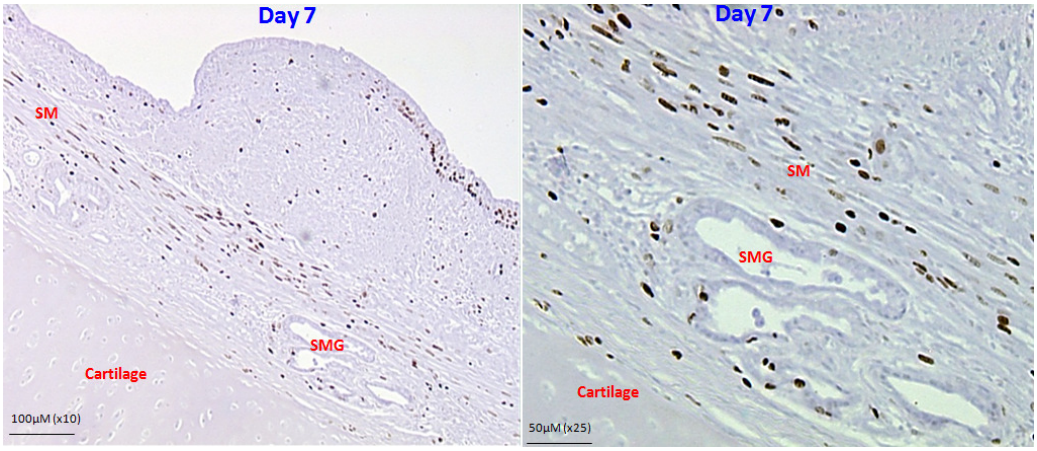


Figure 4.3(e): Airway tissue section at day 7 post injury stained with Ki67 antibody
Whilst most of the damaged site was covered with a new epithelial layer by day 7 proliferative activity could still be discerned, centred mainly underneath the new epithelial layer and on the smooth muscle in the submucosal compartment. The number of Ki67-positive cells was decreased relative to day 3 post injury especially on the epithelium and the submucosal glands.

*** (SMG=submucosal gland; SM=smooth muscle; BV=blood vessels)

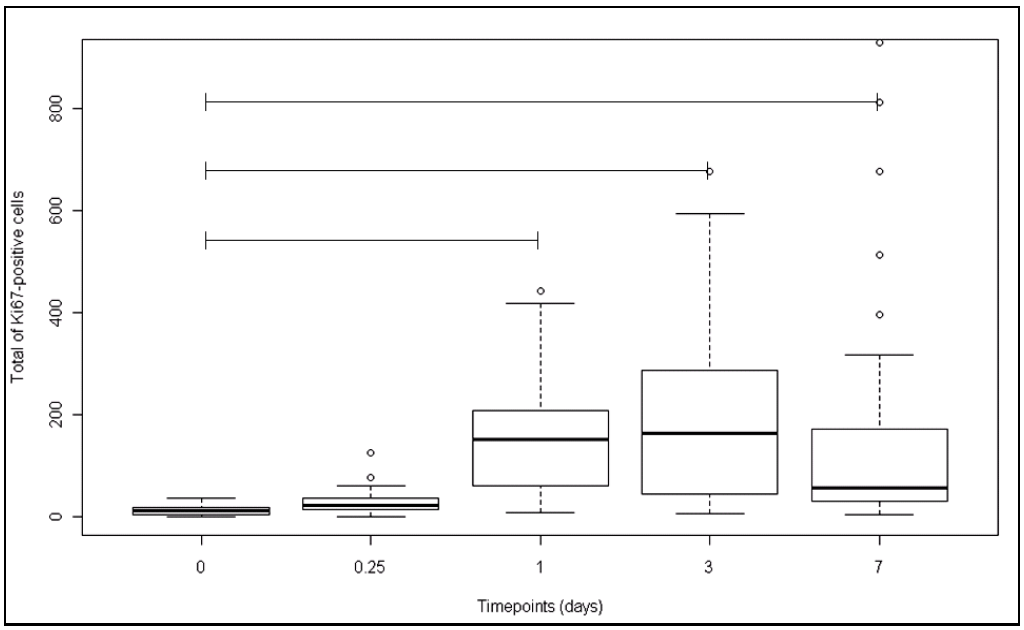


Figure 4.4: The number of Ki67-positive cells between time points with the data for areas of interest and compartments combined.
Although there were proliferative cells observed as early as 6 hours post injury, these cells were not significantly different in number relative to the naïve airway (time point 0). Other time points showed a significant increase relative to numbers observed in naïve airways following injury. Upper and lower box plot margins represent the interquartile range; middle bar indicates the median. The graph was plotted based on the absolute number of cells stained with Ki67 antibody on each area measured. The points outside the ends of the whiskers are outliers.

		MUCOSA			SUBMUCOSA		
		Damaged	**Transitional**	**Undamaged**	**Damaged**	**Transitional**	**Undamaged**
Timepoints (days) (as compared to day 0)	**0.25**	0.839	0.624	0.514	0.237	0.356	0.436
	1	0.063	0.001	0.000	0.000	0.008	0.001
	3	0.001	0.002	0.038	0.000	0.005	0.013
	7	0.003	0.003	0.048	0.000	0.001	0.182

Table 4.3: Table shows the details of *p*-values generated for the distribution of Ki67-positive cells on each areas of interest over the time of injury. The details of its graph are illustrated in the Figure 4.5.

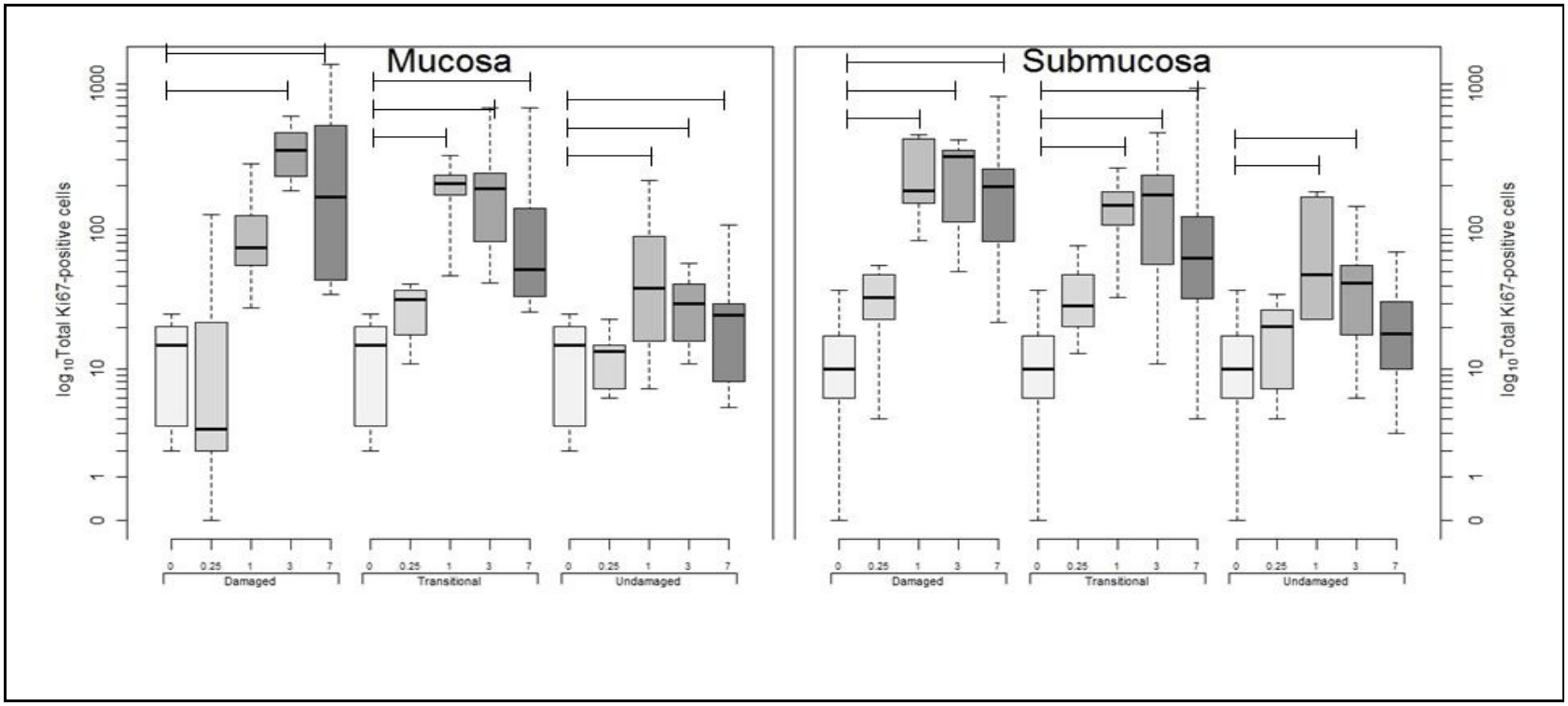


Figure 4.5: Boxplots of Ki67-positive cell distribution arranged to illustrate relationships between compartments (mucosal and submucosal), areas of interest (damaged, transitional and undamaged) and the times of injury.

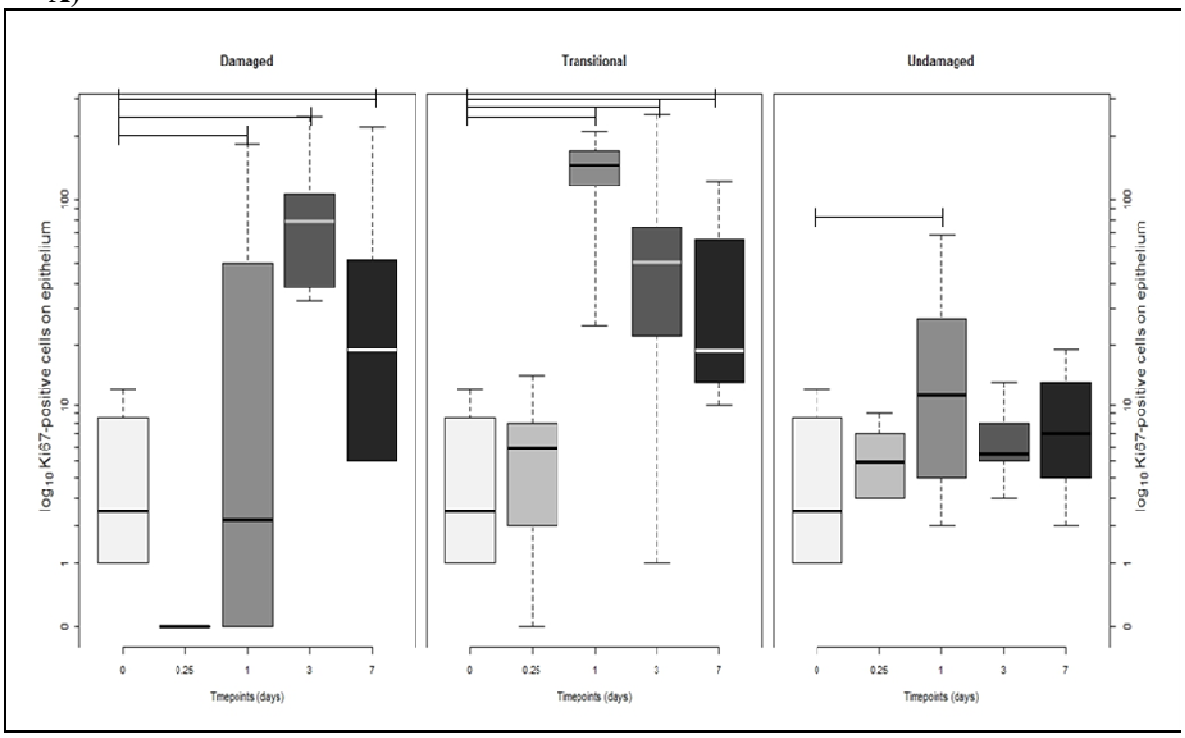
Upper and lower box plot margins represent the interquartile range; middle bar indicates the median. The graph was plotted based on the absolute number of cells stained with Ki67 antibody on each area measured. The points outside the ends of the whiskers are outliers.

With the exception of the proliferative activity evident in the mucosa underlying the damaged area on day 1 and in the submucosa underlying the undamaged area on day 7, the number of proliferating cells increased significantly relative to numbers found in naïve airways from day 1 onwards. The details of the geometric mean values are shown in the Supplement 5 on CD.

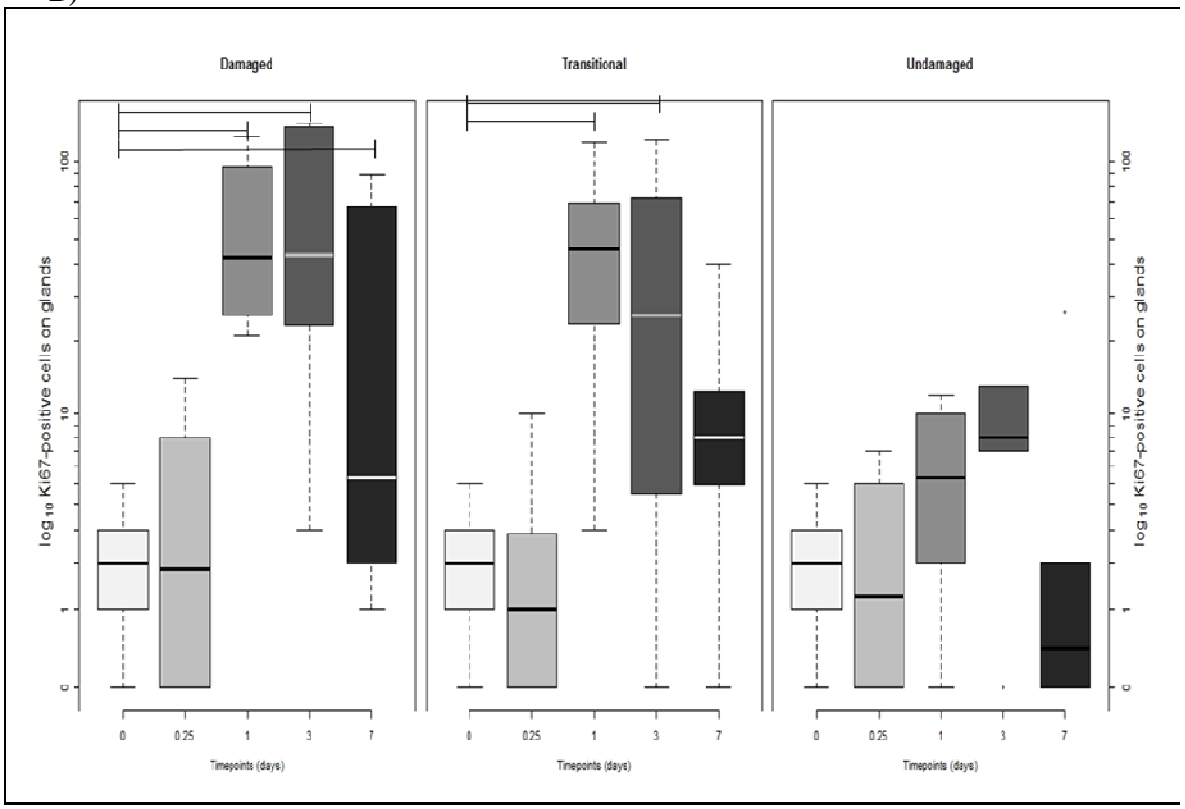
	DAMAGED				TRANSITIONAL				UNDAMAGED			
	Time points (days)				Time points (days)				Time points (days)			
	0.25	**1**	**3**	**7**	**0.25**	**1**	**3**	**7**	**0.25**	**1**	**3**	**7**
EPITHELIUM	0.999	0.001	0.000	0.006	0.861	0.000	0.004	0.005	0.482	0.001	0.725	0.225
BLOOD VESSELS	0.066	0.000	0.000	0.000	0.004	0.000	0.000	0.000	0.428	0.000	0.000	0.002
GLANDS	0.625	0.000	0.001	0.004	0.835	0.004	0.009	0.071	0.659	0.084	0.135	0.256
SMOOTH MUSCLE	0.925	0.464	0.042	0.000	0.968	0.552	0.055	0.025	0.830	0.195	0.683	0.797

Table 4.4: Table shows the *p*-values for the statistical analysis for the number of Ki67-positive cells counted on epithelium, blood vessels, glands and smooth muscle for each areas of interest (damaged, transitional and undamaged) at the various time points following injury as compared to the time point 0. The details of its graph is illustrated in the Figure 4.6. The details of the geometric mean values are shown in the Supplement 5 on CD.

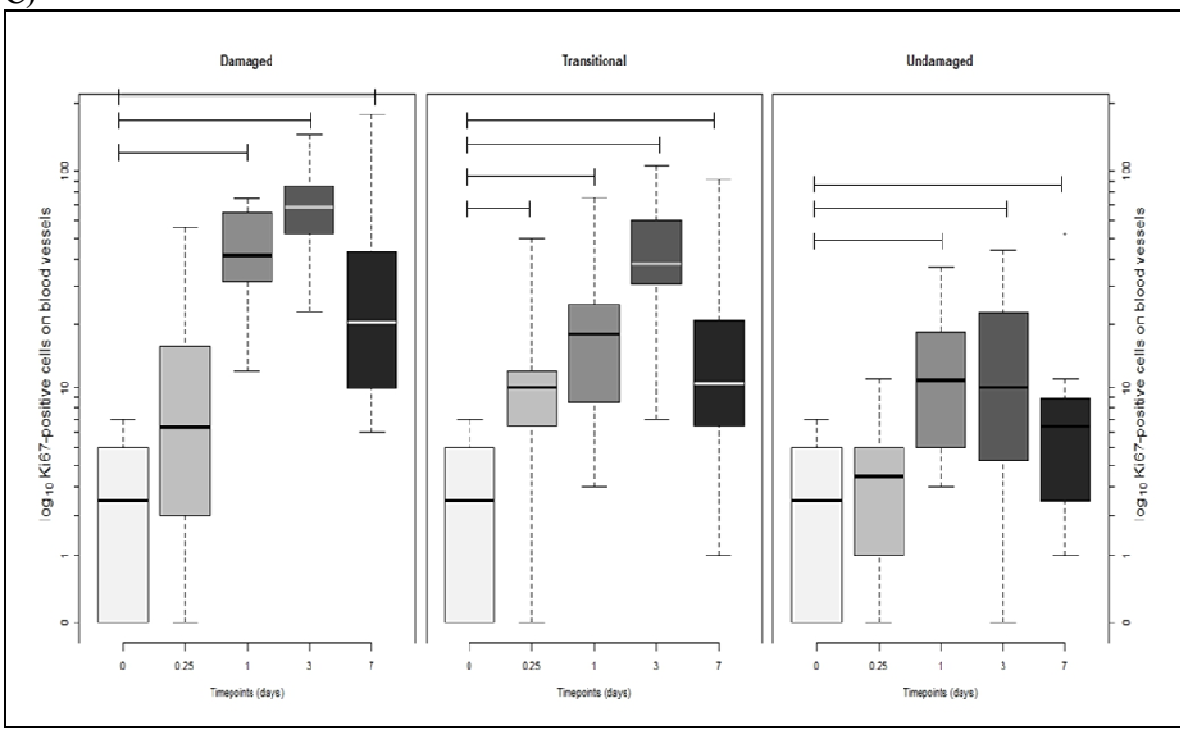
A)


Damaged
Transitional
Undamaged
log10 Ki67-positive cells on epithelium
Timepoints (days)
0
0.25
1
3
7

B)


Damaged
Transitional
Undamaged
log10 Ki67-positive cells on glands
Timepoints (days)
0
0.25
1
3
7

C)


Damaged
Transitional
Undamaged
log10 Ki67-positive cells on blood vessels
Timepoints (days)
0
0.25
1
3
7

D)

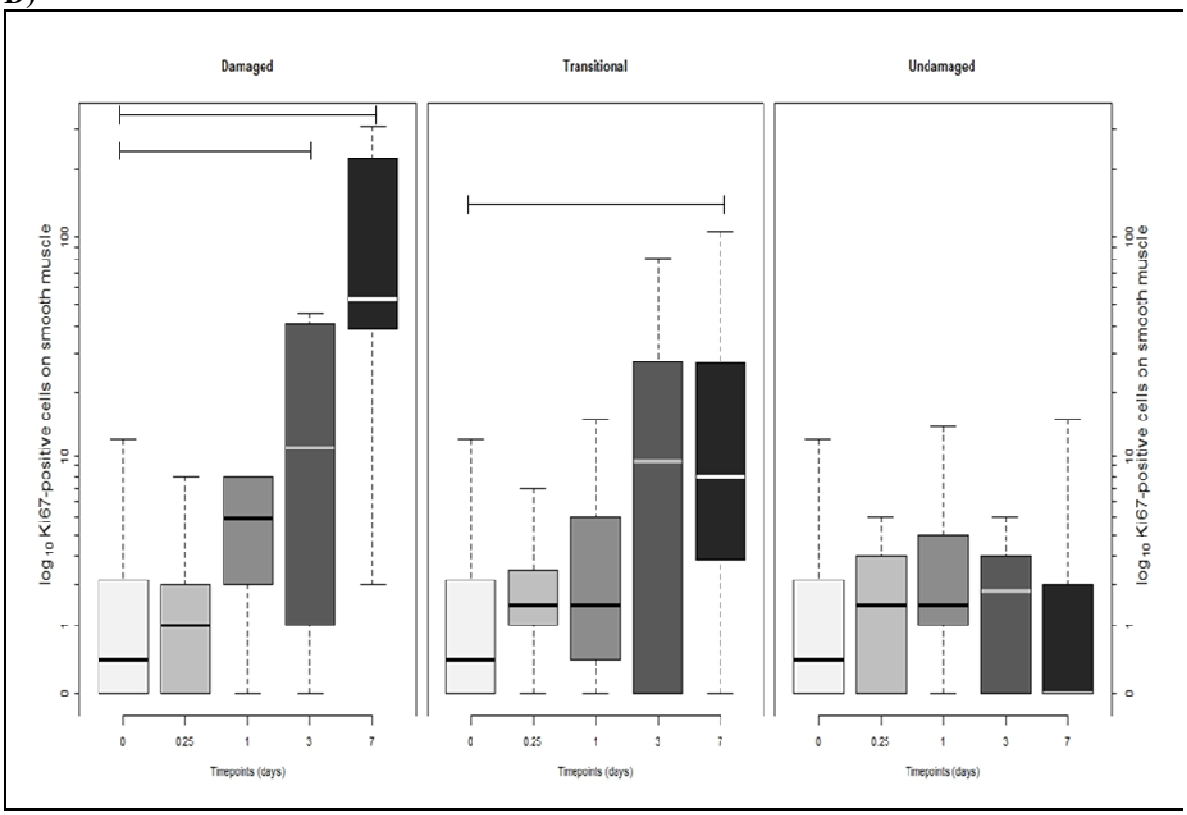

Damaged
Transitional
Undamaged
log10 Ki67-positive cells on smooth muscle
Timepoints (days)
0
0.25
1
3
7

Figure 4.6: The distribution of the Ki67-positive cells on the tissue category.
In all tissues, the number of proliferative cells was increased following injury with significantly higher proliferative activity at the damaged and transitional areas compared to the undamaged areas. Less proliferative activity was observed in the undamaged areas in all tissue types except the epithelium (A), submucosal gland (B), blood vessels (C) and smooth muscle (D). In a smooth muscle tissue shows the proliferative activity is taken place at later time points (day 3 onwards) whilst the epithelium and blood vessels show the numbers of proliferating cells increased significantly ($p<0.05$) in both damaged and transitional areas. In the submucosal glands tissue, cells are significantly active proliferate in both damaged and transitional areas. Upper and lower box plot margins represent the interquartile range; middle bar indicates the median. The graph was plotted based on the absolute number of cells stained with Ki67 antibody on each area measured. The points outside the ends of the whiskers are outliers. The details of the geometric mean values are shown in the Supplement 5 on CD.

4.4.6 AB-PAS staining

The effect of the brushing on the mucus-producing cells within the epithelium was similar to that described in the Chapter 2; Section 2.4.2.3. Two tissue categories contain mucus-producing cells, the epithelium (goblet cells) and the submucosal gland (serous cells). AB-PAS-positive cells were easily identified on the epithelium and glands of the normal airway (Figure 4.7 (a)). The number of PAS-positive cells in total was decreased in damaged and transitional areas as compared to the undamaged areas ($p<0.05$) in both tissues. The morphology of the mucus-producing cells (particularly the goblet cells (GC)) on the epithelium at the transitional area, changes significantly over the time of injury [(Figure 4.7 (b)) to (Figure 4.7(c))]. This cell was generally absent on the damaged epithelium and gradually began to reappear by 3 days post injury (Figure 4.7(d)). At day 7 post injury, greater numbers of GCs were observed at the damaged site although the morphology of these cells was dissimilar to that seen on the epithelium of the undamaged area (Figure 4.7 (e)) compared to the normal morpholopgy of the GCs as shown in the Figure 4.7(a).

Table 4.5 shows the details of p-values generated from statistical analysis to measure the distribution of AB-PAS positive cells on the tissues of interest (epithelium and submucosal gland) over the various times following injury as compared to the time point 0. In the epithelium (Figure 4.8(A), as compared to the naive airway (time point 0), in the damaged and transitional areas, the number of PAS-positive stained cells was significantly reduced at all time points ($p<0.05$) following injury (except at days 3 and 7 in the transitional area) whilst in the undamaged area, the number of PAS-positive cells was significantly reduced at day 3 post injury ($p<0.05$). In the submucosal gland tissues (Figure 4.8(B), the number of cells stained with AB-PAS staining in both damaged and transitional areas was significantly less at 6 hours and day 1 post injury compared to naïve (time point 0). In the undamaged area, there was no significant difference observed in any time points as compared to naïve sites. The details of the R outputs are shown in the Supplement 6 on CD.

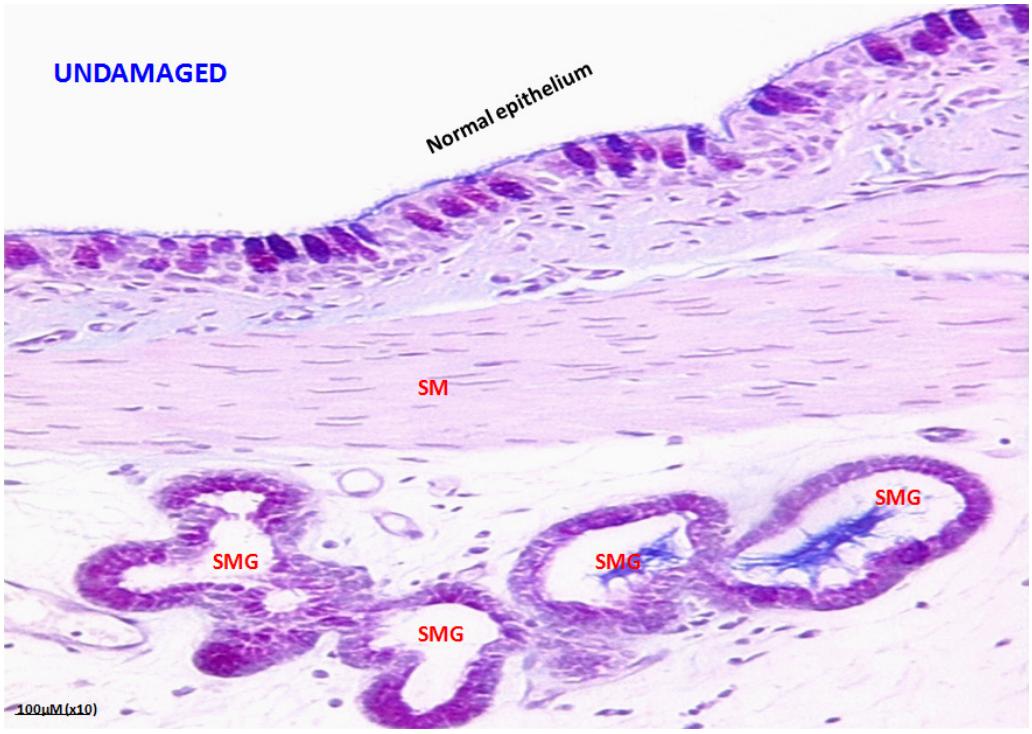


Figure 4.7(a): AB-PAS staining of an undamaged airway
The figure illustrates the normal morphology of the goblet cells in the epithelium and serous cells in the SMGs (AB-PAS staining). (* SMG = submucosal gland; SM = smooth muscle)

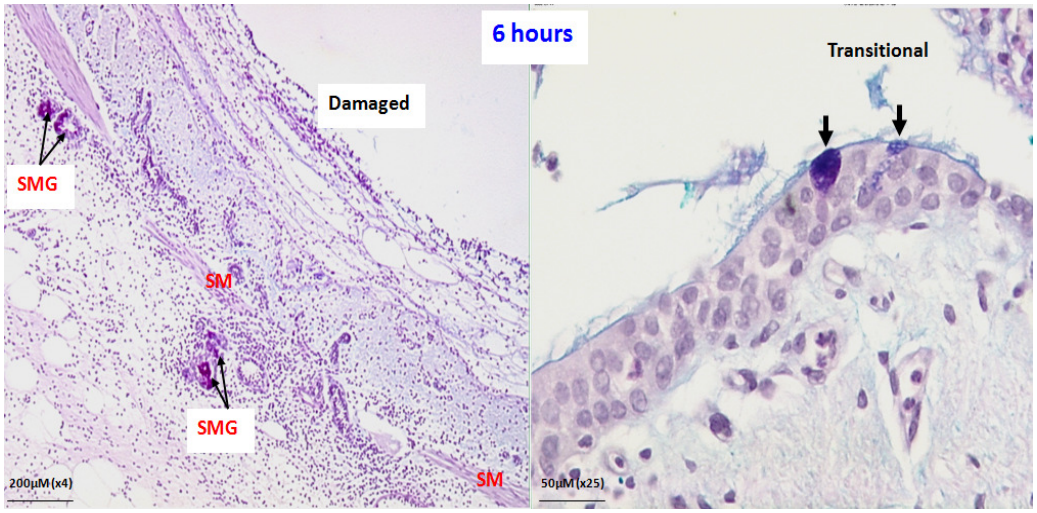


Figure 4.7(b): AB-PAS staining of an airway section at 6 hours post injury
The figure illustrates the absence of GCs at the lesion site and the morphology of this cell at the transitional area is changed (picture insert) where PAS staining was found at the apex of the epithelial cells (arrows). (* SMG = submucosal gland)

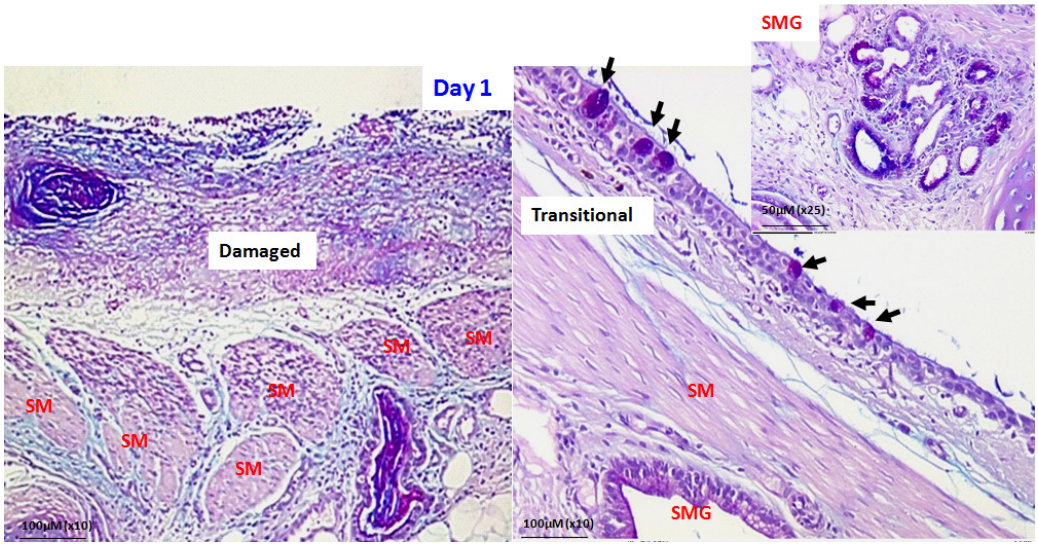


Figure 4.7(c): AB-PAS staining of an airway section at day 1 post injury
At this time point a clearer dedifferentiation of the GCs at the transitional area can be observed (arrows) and intense PAS staining can be discerned at the submucosal glands (picture insert). At the damaged site, there was still evidence of mucosal loss at this time point and centrally within the damaged area GCs loss was still evident. (* SMG = submucosal gland; SM = smooth muscle)

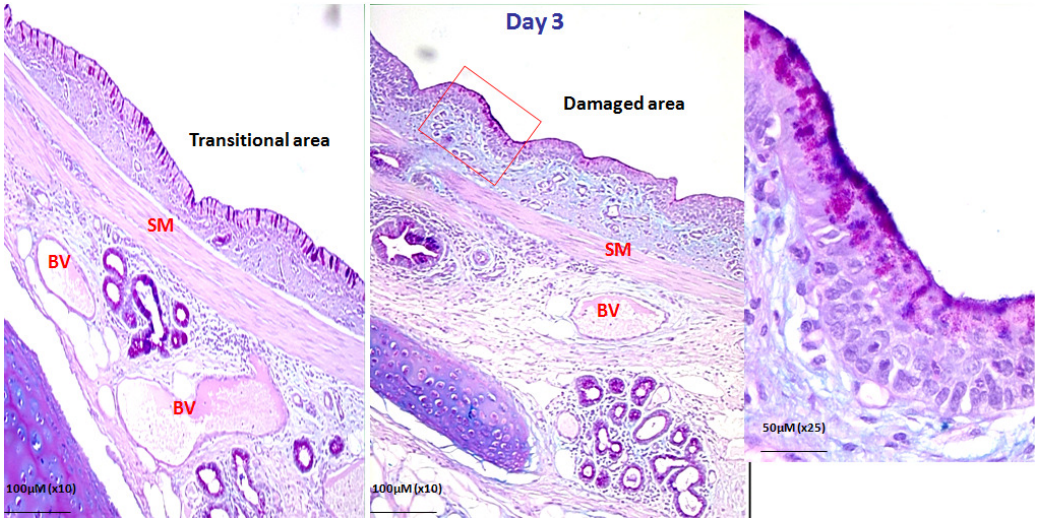
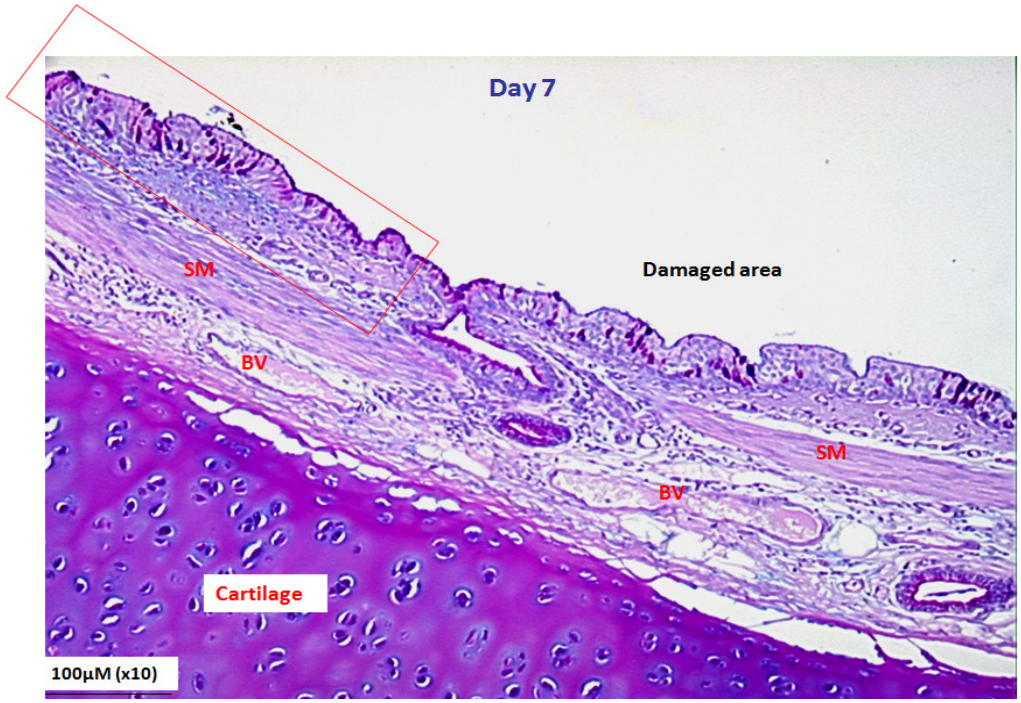


Figure 4.7(d): AB-PAS staining of an airway section at day 3 post injury
At the brushed site (right panel), there were some thickened areas of epithelial hyperplasia (insert) and the AB-PAS staining indicated that the goblet cell numbers increased compared to day 1 especially at the transitional area (right panel). (* SM = smooth muscle; BV = blood vessel)

A)



B)

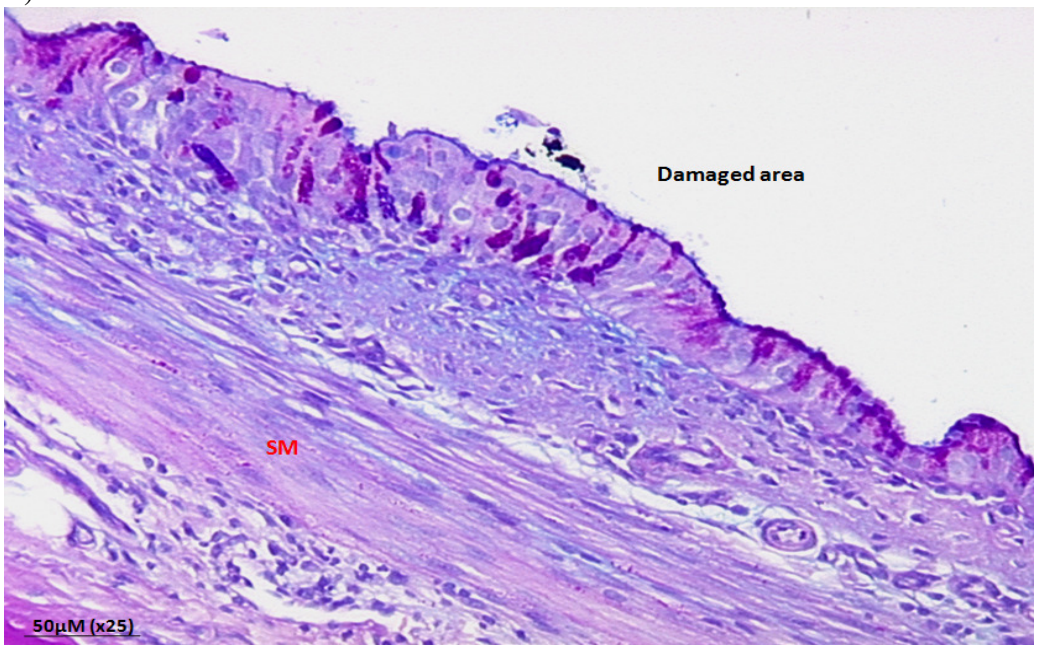
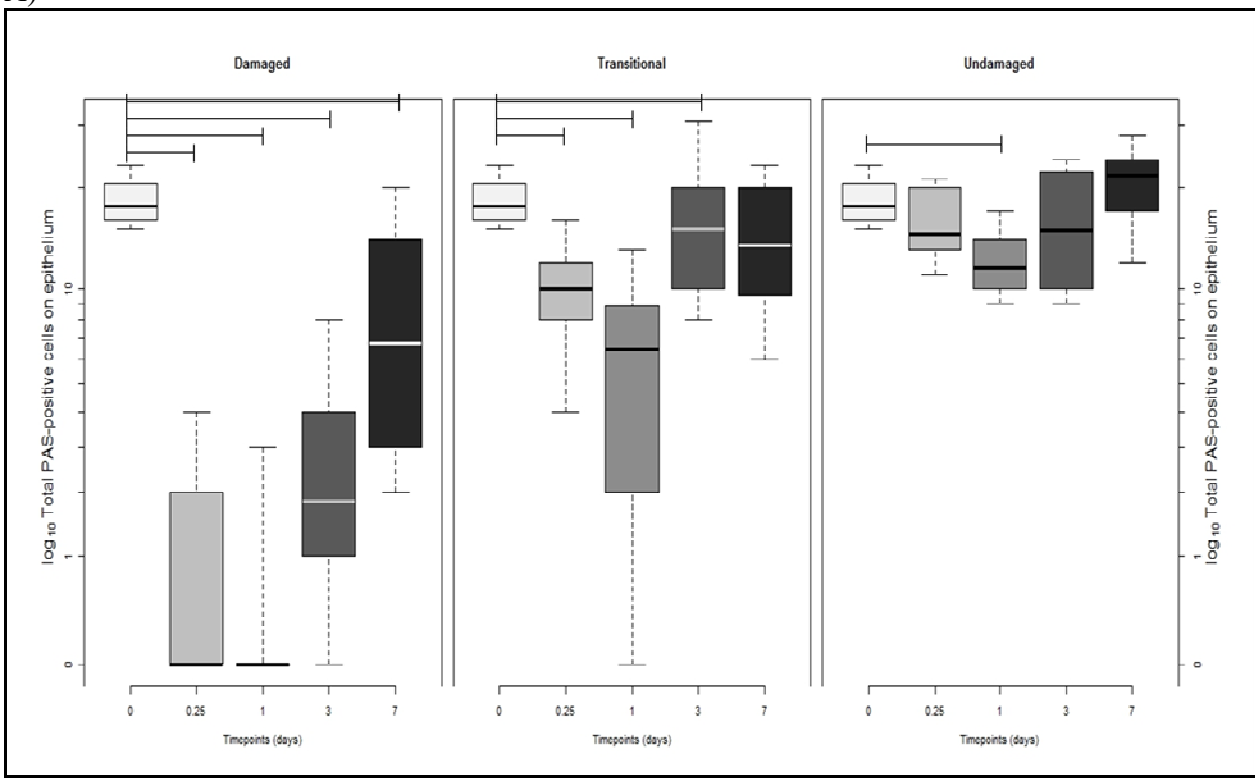


Figure 4.7(e): AB-PAS staining of an airway section at day 7 post injury
At this time point, the epithelium at the brushing site was intact and some areas were lined with stratified squamous epithelium (not depicted). The number of the goblet cells had increased at the damaged (A) area at this point in time and the majority of these newly regenerated PAS-positive stained epithelial cells showed morphology inconsistent with GCs (B – higher magnification of the figure A) seen in the undamaged area (Figure 4.7(a)).

	DAMAGED				TRANSITIONAL				UNDAMAGED			
	Time points (days)				**Time points (days)**				**Time points (days)**			
	0.25	**1**	**3**	**7**	**0.25**	**1**	**3**	**7**	**0.25**	**1**	**3**	**7**
Epithelium	0.005	0.006	0.000	0.023	0.001	0.000	0.020	0.079	0.234	0.010	0.265	0.592
Glands	0.130	0.022	0.015	0.644	0.009	0.016	0.443	0.838	0.845	0.562	0.733	0.860

Table 4.5: Table shows the *p*-values for the statistical analysis for the number of AB-PAS -positive cells counted on epithelium and submucosal glands for each areas of interest (damaged, transitional and undamaged) at the various time points following injury as compared to the time point 0. The details of its graph is illustrated in the Figure 4.8. The details of the geometric mean values are shown in the appendix. The details of the geometric mean values are shown in the Supplement 6 on CD.

A)


Damaged
Transitional
Undamaged
log10 Total PAS-positive cells on epithelium
0
1
10
0
0.25
1
3
7
Timepoints (days)

B)

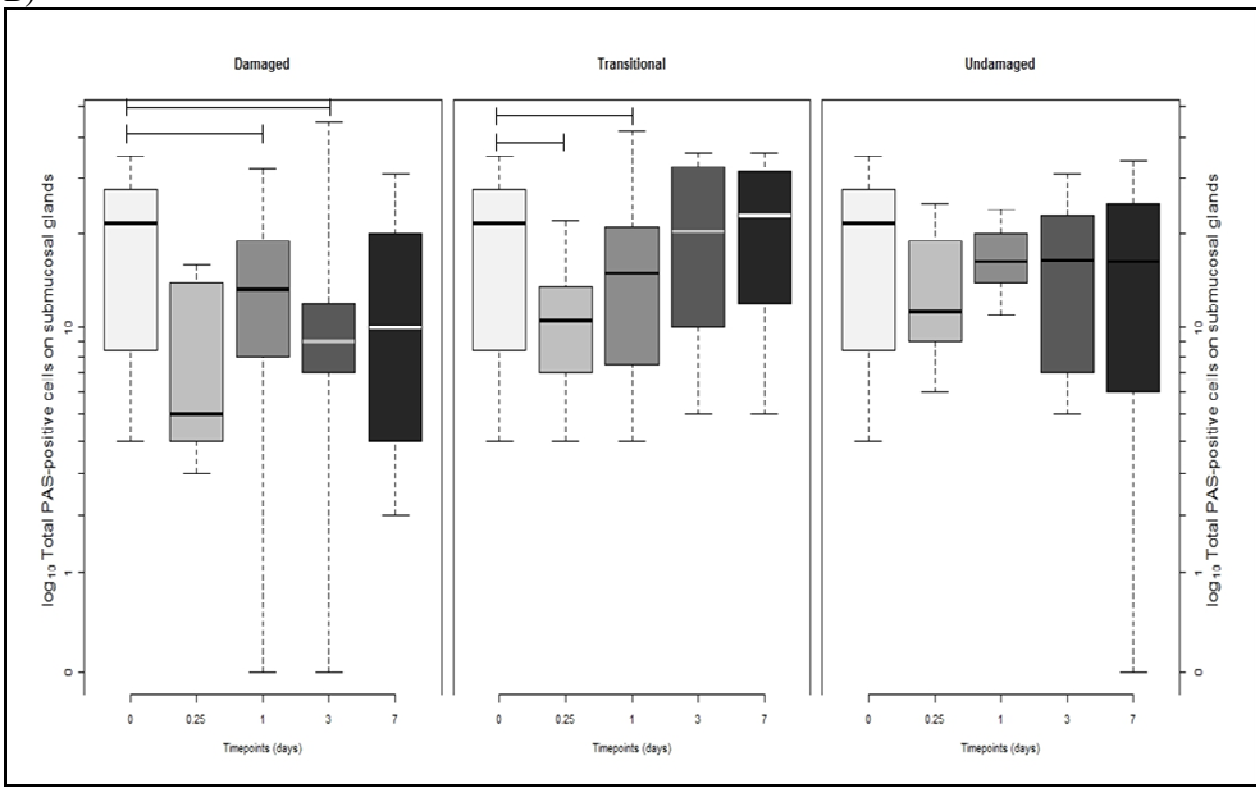

Damaged
Transitional
Undamaged
log10 Total PAS-positive cells on submucosal glands
0
1
10
0
0.25
1
3
7
Timepoints (days)

Figure 4.8: The number of the PAS-positive stained cells within the epithelium and submucosal gland structures of the airway wall.
The number of positive-stained cells on the epithelium in the damaged area was significantly reduced at all time points relative to numbers seen in naïve airways. A similar phenomenon was apparent in the transitional area (with the exception of the reduction at day 7 which was not significant), whilst only at day 1 was there a significant reduction in the numbers of PAS-positive cells within the undamaged area relative to naïve airways. In the submucosal gland tissue, the number of PAS-positive cells was significantly reduced in the damaged area relative to naïve airways at days 1 and 3 post injury. Significantly reduced numbers of PAS-positive cells relative to naïve airways were also identified within the transitional area at 6 hours and day 1 post injury. Upper and lower box plot margins represent the interquartile range; middle bar indicates the median. The graph was plotted based on the absolute number grid points landing on each cell stained with AB-PAS staining on each area measured. The points outside the ends of the whiskers are outliers. The details of the geometric mean values are shown in the Supplement 6 on CD.

4.5 DISCUSSION

The main goal of this study was to confirm the characteristic histopathological changes that are found in the airway epithelium in response to bronchial brushing injury. This study followed on from our pilot studies using *in vivo* BrdU labelling of airway epithelial cells (Chapter 2).

This study was designed such that two sheep were subjected to brushing at three time points in each of four different combinations of time points (protocols). Two segments were subjected to brushing at each time point (Table 4.1). One site per time point was harvested for subsequent microarray analysis and the other for analysis of the histopathological changes of the airway following injury. As a consequence the comparison of molecular and histopathological changes at each time point was made between different segments of the lung. As in the previous chapter (Chapter 2) effort was directed towards confirming that airway structure (in terms of the proportion of epithelium, glands, smooth muscle, blood vessels and cartilage) was consistent between separate lung segments. The study confirmed that there was no segment effect on the proportion of these airway tissues within the airway wall regardless of their position within the tracheolbronchial tree and suggests that injury sites within different segments are structurally comparable.

The experimental design was adapted from the previous preliminary experimental approach involving two different protocols (Chapter 2; Protocols A and B). The proliferative response of the BrdU-positive cells observed in Chapter 2 suggested no difference in terms of histological changes and cellular (proliferative) activity in both protocols, at least at the histological level. Therefore, this study was adapted from one of these protocols (Protocol A) and brushing employed as described in Chapter 2. Tissue from a naïve segment was collected from each sheep (n=8) as a control based on the assumption that minimal cellular and molecular responses would occur at such naïve sites. This approach was based on our observations in Chapter 2 where there was scant evidence of histological changes or BrdU incorporation in undamaged sites. Similarly, in

this study, from the histopathological point of view (HE, AB-PAS, Ki67), there were no obvious changes in terms of airway morphology, goblet cell morphology or proliferation in the naïve regions in agreement with the previous study.

At 6 hours post injury, there was disruption of the airway mucosa including the basement membrane similar to that observed in the Chapter 2. The single layer of squamous epithelial cells migrating from the border of the lesion is thought to be an early stage of the airway repair. The histopathological features of the airways at day 1 differed to those observed at day 3 post injury. At day 1, the damaged regions were clearly visible and the transitional area comprised single- to multiple layers of epithelial cells. At day 3, although within some areas of damage there was still evidence of mucosal haemorrhage, at the transitional area there was evident of thickened epithelium comprising multiple layers of undifferentiated cells.

In general, based on the histopathological changes observed in the airway sections, the surface of the injured site was fully covered, with at least a partially differentiated airway epithelium by 7 days post injury. This finding was similar to that previously described in study of bronchial brushing in healthy human volunteers (Heguy A et al., 2007) which revealed that at day 7 post injury, the damaged airway was completely covered by epithelial cells, the majority of which appeared to be undifferentiated. At day 7 post injury, peripheral areas of the injured epithelium had regained a normal morphology likely due to the fact that the airway repair process was initiated from the transition region around the lesion and progressed towards the central area. Therefore the epithelium at the central area of damage was still only partially differentiated by day 7. In the transition region, the proportion of the airway epithelium was not different at day 7 to that seen in the undamaged (day 0) area thus indicating that the airway epithelium at the transitional area had been fully covered (most areas had begun to differentiate) with the new epithelium following injury. This finding was supported by the analysis of the AB-PAS staining on the epithelium where the total number of PAS-positive cells had returned to the level of the naive airway (time point 0). The number of

PAS-positive stained cells in the damaged region was also increased at day 7 compared to the earlier time points ie it had still not returned to normal. This finding supports the concept that at this time point the newly regenerated epithelial cells had begun to differentiate into specialised epithelial cells (goblet cells in particular). Although the number of PAS-stained cells was increasing towards normal at this time point (day 7), many of them had an unusual morphology (Figure 4.7 (e)). The preliminary study (Chapter 2) suggested that the central part of the damaged area had still not regained a fully differentiated epithelium even at day 14 post injury and numbers of AB-PAS stained cells were similar to that seen in day 7 post injury.

In relation to assess the proliferative activity of the airway epithelial cells in response to bronchial brushing, this study used Ki67 marker instead of *in vivo* labelling using BrdU. The advantage of Ki67 over BrdU in functioning as a proliferative marker is that the Ki67 labels all proliferating cells at all stages of the cell cycle except during the resting phase (G_0). Variation in staining intensity likely reflects staining during different cell cycle phases (Gerdes J et al., 1984). The increase in proliferative activity within the mucosal compartment may reflect the process of replacing the damaged epithelium. Increased proliferative activity in the mucosal compartment was not only observed at the damaged site but also at the transitional areas bordering the lesion. The submucosal gland ducts have previously been described as the location of stem cell niches from which epithelial cells migrate to cover the denuded areas (Borthwick DW et al., 2001). In the present study, the proliferative activity in the submucosal gland was not only increased in the damaged region, but also in the transitional region bordering the lesion. This finding supports the initial hypothesis that submucosal gland as a source of progenitor cells (stem cell niche). The number of Ki67 positive cells peaked at days 1 and 3 post injury within the transitional area (Figure 4.6(A); epithelium). This finding suggests that regeneration and repair processes were actively undertaken at these time points.

The delay in the peak of proliferation observed in this study (as compared to the small animal model system described in Chapter 2) might be related to the severity of the brushing protocol. In our case the airway mucosal epithelium was completely removed (including the basement membrane and elements of the smooth muscle layer) as discussed in the Chapter 2 (Chapter 2: Section 2.5; Discussion). Indeed, when the analysis was performed in order to assess the effect of the brushing on the tissues in mucosal and submucosal compartments, it was revealed that not only was the proportion of the epithelium significantly different to naïve ($p<0.05$) with no epithelium at the damaged site at the earlier time points (6 hours and day 1) but the proportion of the smooth muscle was reduced at the earlier time points. By day 7 post injury the proportion of the smooth muscle tissue had returned to that of naïve airways. This regeneration of the smooth muscle appeared to be happening between 3 and 7 days post injury. A corresponding increase in the number of Ki67-positive cells in the smooth muscle tissue at the damaged and transition regions at days 3 and 7 was observed (Figure 4.2; smooth muscle). Although the brushing had no significant ($p>0.05$) effect on the proportion of the submucosal gland tissues, the distribution of the Ki67-positive cells on this tissue especially underneath the damaged region might contribute to the overall proliferative activity observed on the damaged area at the mucosal compartment as to compensate for brushing-induced epithelium loss.

Overall the proportion of blood vessels did not significantly alter following the brushing procedure at any of the time points, however there was a significant increase in the proportion of the blood vessels in the damaged region compared to undamaged at day 3 post injury which was likely due to the visible angiogenesis, a normal feature of injury and repair processes. The increased number of proliferating cells on the blood vessels in all areas of interest as compared to naive airway suggests that some of the proliferative activity observed in this study might have its origins in blood-derived cells. This finding contrasts with our initial assumption in the Chapter 2 in which proliferative activity was not necessarily adjudged to be of blood cells origin. Dual staining would help to confirm this fact. However, the lack of appropriate reagents constrains attempts to follow this

through. There is no doubt that blood-derived cells contribute to airway epithelial regeneration and repair (Abe S et al., 2004, Asakura A and Rudnicki MA, 2002) by migrating towards the damaged site but whether or not these cells directly proliferate and differentiate to become specialised epithelial cells could not be ascertained in this study.

Consistent with the pathophysiological changes observed at day 7 post injury (with evidence of undifferentiated airway epithelial cells within the central area of the damaged site), a proliferative activity measured by the distribution of Ki67-positive cells was highly significant compared to the undamaged area. Although the total numbers of Ki67-positive cells were found to be lower compared to day 3 post injury this finding might suggest that the majority of the proliferating cells were involved in the late stage of the cell cycle (G_2 and M) with relatively few still at the early stage of the cell cycle phase. This observation is based on the various intensities of the Ki67 staining on proliferating cells with dense staining indicate cells at the late stage of the cell cycle and faint staining indicating cells at an earlier cell cycle phase (Figure 4.3). In a study of airway repair in healthy volunteers, Heguy et al., (2007) showed that most of the genes up-regulated on microarray analysis at day 7 post injury were those involved in the late stage of the cell cycle, including those governing signal transduction and metabolism (Heguy A et al., 2007).

Mucus producing cells were involved in the early process of dedifferentiation involving the airway epithelial cells bordering the lesion either by changing their shapes and morphology to migrate to the denuded area or through the intermediate cells (basal or para-basal cells) migrating towards the denuded area to become squamous epithelial cells that would eventually cover the brushed area. The AB-PAS staining revealed that there was a significant reduction of AB-PAS-positive cells at 6 hours, days 1 and 3 post injury with consistent loss of epithelium at the earlier time points (6 hours and day 1) following bronchial brushing (Figure 4.7(b) and (c)). The morphology of the mucus-producing cells at the transitional areas changed remarkably compared to the staining on

these cells within undamaged areas. Although there was an evidence of regeneration of epithelial cells or mucus-producing cells (goblet cells in particular) at day 3 post injury (Figure 4.7(d)), the total number of cells stained for AB-PAS was still less than seen in undamaged areas.

A finding consistent with that demonstrated in Chapter 2 was the intact mucus 'scaffold' overlying the damaged area at day 3 post injury. The AB-PAS staining was arranged in a thin covering over the top of the injured epithelial surface, and this layer was packed with neutrophils. Whether this consistent finding suggests a key step in the process of epithelial repair and regeneration remains to be seen.

4.6 CONCLUSION OF THE CHAPTER

In conclusion, this study demonstrates a novel model for the study of airway injury and repair via bronchial brushing in large animal lung. The study confirmed the findings described in the Chapter 2 in terms of the pathophysiological changes of the airways over time from injury involving dedifferentiation, proliferation and re-differentiation to restore almost completely regenerated airway wall morphology by seven days. Indeed, although the brushing technique used in this study perturbed both mucosal and submucosal morphology, the ability of the airway to regenerate a new epithelium and other airway components such as smooth muscle and glands as early as 3 to 7 days post injury was remarkable. Consistent with the findings in the Chapter 2, the physical lung injury model carried out in this study appears to show similar features to the mechanical injury techniques previously described in small animal model systems. It is anticipated that subsequent high throughput molecular analysis will provide opportunities to associate the histopathological findings observed in this chapter with the molecular changes in response to bronchial brushing. Thus this approach will lead to new insight into the functional genomics of sheep airway injury and could provide information relevant to the pathophysiology of airway disease.

CHAPTER 5

TRANSCRIPTIONAL PROFILING OF OVINE AIRWAY REGENERATION AND REPAIR FOLLOWING BRONCHIAL BRUSHING

5.1 INTRODUCTION

A microarray study is a high throughput expression profiling method which requires high quality and integrity of RNA samples to ensure reliability, accuracy and reproducibility of the microarray results. Expression profiling studies have been widely reported since the advent of this technology in the early 1990's and such use has extended to studies of lung development (Mariani TJ et al., 2002), lung disease pathogenesis (Hansel NN and Diette GB, 2007, Wang IM et al., 2008, Raman T et al., 2009), and lung repair following either chemical or physical injury (Brass DM et al., 2008, Heguy A et al., 2007, Jeyaseelan S et al., 2004, Shultz MA et al., 2004, Yoneda K et al., 2003). Naturally these, and other, studies have encompassed a wide range of complex experimental design to meet defined objectives and obtain a better understanding of the complex molecular structure that underlies lung biology.

During the development of this thesis, preliminary studies had been carried out involving studies of the pathophysiological (Chapter 2) and molecular aspects (Chapter 3) of airway repair. Realising the fact that both these studies are limited in terms of the numbers of sheep used and the low number of genes studied, a more robust study was designed to achieve a clearer understanding of cellular and molecular changes during airway regeneration and repair and how these changes might relate to the pathophysiological changes described in Chapter 4. An advantage of high-throughput microarray technologies is that they can offer an opportunity to gain insight into the global gene expression profiles of disease processes such as asthma, leading to the identification of new asthma-associated genes and pathways (Tölgyesi G et al., 2009).

Microarrays have provided novel insight into many complex human diseases and microarray-based discovery can be classified into three components, biomarker detection, disease (sub) classification and identification of causal mechanisms (Bhattacharya S and Mariani TJ, 2009). The application of transcriptomic analysis of genes involved in airway epithelial repair has mostly been performed in samples taken from patients with lung-related diseases such as asthma and COPD. The use of lung

tissue is ideal for the investigation of asthma pathogenesis, and several recent studies have analyzed global gene expression in bronchial biopsy specimens (Hansel NN and Diette GB, 2007). In COPD, studies have mapped comparative gene expression profiles of lung tissues from patients with different stages of COPD relative to healthy smokers or non-smokers. Such studies have revealed a number of differentially-regulated genes associated with COPD progression, which include genes involved in the regulation of inflammation, extracellular matrix, cytokines, chemokines, apoptosis, and stress responses (Chen ZH et al., 2008). Endoscopic bronchial brushing procedures for airway epithelial cell sampling are more effective in sampling specific target areas within the lung.

However, in the context of understanding the mechanisms involved in airway epithelium regeneration and repair in the normal condition (healthy individual with no lung-related disease), to the best of our knowledge, only one study has applied this combination of bronchoscopy and mechanical brushing of small regions of the airway epithelium of normal individuals, to force the epithelium to undergo regeneration *in vivo* (Heguy A et al., 2007). A limitation of this study was that the expression profiles at brushed sites were compared to the unbrushed airways at only two time points (Days 7 and 14 post-injury). Therefore, this study provides no insight into the early molecular changes following injury. However, at day 7 compared with resting epithelium, there were substantial differences in gene expression pattern, with a distinctive airway epithelial "repair transcriptome" of actively proliferating cells in the process of re-differentiation (Heguy A et al., 2007).

This study was designed to build on the observations from Chapters 2 to 4 and provides a more comprehensive knowledge of the genes and pathways involved in airway epithelial repair. The recent availability of an ovine microarray chip (Agilent Technology) provided an opportunity to investigate a time-course of molecular changes in our sheep model of bronchial brushing induced airway injury from 6 hours to 7 days, and compare those expression profiles with uninjured (naive) airways.

As reviewed by Murphy D, (2002), the basic concept of microarray analysis involves labelling the RNA extracted from the tissue of interest to generate the *target*—the free nucleic acid sample whose identity or abundance is being detected. This is hybridized to the tethered *probe* DNA sequences corresponding to specific genes that have been affixed, in a known configuration, onto a solid matrix. Hybridization, based on Watson and Crick base pairing, between probe and target provides a quantitative measure of the abundance of a particular sequence in the target population. This information is captured digitally and subjected to various analyses to extract biological information. Comparison of hybridization patterns enables the identification of mRNAs that differ in abundance in two or more target samples (Murphy D, 2002).

There are two common approaches to microarray studies, depending on the objective of the designed study. A one-colour (intensity-based) approach can be used to compare treated and untreated animals exposed to certain chemicals (Brass DM et al., 2008, Yoneda K et al., 2003) whereas a two-colour (ratio-based) approach can be used to assess the expression differences between mRNA isolated from two types of tissues such as parenchyma and airway (Shultz MA et al., 2004). The addition of a one-colour platform to the Agilent gene expression portfolio gives the ability to take advantage of the enhanced performance and sensitivity of Agilent's 60-mer oligonucleotide microarrays (Xiao J et al., 2006). When an intensity-based microarray platform is used for gene expression analysis, each sample is hybridized on a single microarray therefore, a one-colour solution provides the ability to compare the measured gene expression output of a microarray directly across other microarrays to generate new and multiple ratiometric measurements (Xiao J et al., 2006). This differs from a two-colour approach where all gene expression ratios are generated only from two samples compared on the same microarray. Two-colour microarray experiments typically use cyanine 3 and cyanine 5 dyes. In one-colour experiments, the cyanine 3 dye is used because it is less susceptible to degradation by environmental factors such as ozone, pH, and organic solvents, compared with cyanine 5 dye.

The nature of such high-throughput platforms places considerable technological demands that, if compromised, can confound the fidelity of the results. Such demands range from the precision and reproducibility of the print heads in spotting the arrays to the uniformity of the hybridisation process and the stringency and consistency of the slide washing process. The effects of compromise in such technology are usually manifest in systematic bias that appears irrespective of the underlying biological variation present within and between arrays on slides. One type of problem is regional bias, in which one region of a chip shows artifactually high or low intensities (or ratios in a two-channel array) relative to the majority of the chip (Reimers M and Weinstein JN, 2005). Factors such as temperature, liquid flow rate or RNA diffusion rate may differ among different regions on the array. Higher local off-spot background near the edges is normally due to less thorough washing near the edge of the slides which push up the spot measures, although not always predictably. Another problem is due to a bubble which remains after filling an Affymetrix cassette. This bubble will not travel uniformly over the chip during hybridization mixing and may get stuck or move in an irregular circuit. Scratches and other manufacturing imperfections can make a difference (Reimers M and Weinstein JN, 2005).

The presence of such bias can be detected, and quantified by various quality-control assessments that are applied to the scanned images from each array and area feature of the output from commercial platforms such as from Agilent arrays.

To counteract and allow for the effects of such bias array experiments will often include some process of normalization. Common methods of normalization applied to two-colour arrays include global normalization (where a measure of distribution (eg mean, median or percentile) of different arrays is adjusted to a constant, and locally weighted scatterplot smoothing (LOWESS), where a locally weighted polynomial regression of the intensity scatterplot is used in order to obtain the a calibration factor. Similar global normalization strategies such as percentile normalization, dividing the expression of measured genes by the intensity of pre-selected housekeeping genes, dividing expression

values of genes of interest by the sum of all signals or finding the most probable consistent pairwise scaling based on central tendency of expression ratios can be applied to single-colour arrays. Importantly however, as remarked by Kreil and Russell (2005), “there is no silver bullet – no single ‘catchall’ method solves the problem of appropriate microarray data analysis” and, importantly, the normalisation method used should be justifiable in the particular biological context of the experiment and not just performed "as standard". Normalization should not be performed for the sake of normalization, but in response to demonstrable systematic bias attributable to the process itself (Kreil DP and Russell RR, 2005).

Protocols for collection and storage of lung tissue samples for subsequent extraction of RNA with the required high quality and integrity for microarray analysis have been reported (Chomczynski P, 1992, Keating DT et al., 2008). RNAlater, a proprietary tissue storage reagent that rapidly permeates tissue to stabilize and protect cellular RNA *in situ* in unfrozen specimens, is a popular medium to preserve samples for subsequent molecular analysis such as microarray studies. A study by Keating DT et al., (2008) suggested that storage in RNAlater preserved RNA quality to a level appropriate for gene expression studies (Keating DT et al., 2008). This study contrasted with the experience of Micke P et al., (2006) who suggested that preserving the tissue samples in RNAlater introduced variability in measured gene expression levels. As an alternative approach, flash freezing in liquid nitrogen is quick and effective in preserving tissue. Indeed a recent study demonstrated that RNA is relatively well preserved when fresh tissue is rapidly frozen for tissue bio-banking with expression levels of selected genes remaining stable in such non-fixed tissue specimens for a long period of time (Micke P et al., 2006).

We have established that the semi-quantitative RT-PCR analysis in Chapter 3 revealed substantial variation in terms of the observed expression profiles of the selected genes. Whilst the cause of this variation is unknown, RNA quality and integrity may be contributing factors. Expression profiling the time course of corresponding molecular

changes during the repair process requires the preparation of high quality RNA from the injury sites. These airway tissue samples are relatively small and cartilaginous which, at least on an anecdotal level (R Talbot, personal communication), can lead to problems in preparing RNA of the concentration and quality required for microarray analysis. A related concern was that the tissues under study may contain an increased proportion of apoptotic cells containing degraded RNA which could contribute to problems with quality (Aigner T and Kim HA, 2002). Regarding the amount of RNA required for microarray analysis, Wang S et al., (2003) suggested that 5µg of total RNA is the minimum that should be used for direct labelling procedure. While lower amounts of RNA input may be sufficient to yield detectable and accurate log ratios for some transcripts, less than 2 µg total RNA may result in lowered detection of less abundant transcripts and ultimately yield reduced amounts of data (Wang S et al., 2003).

There are various fundamental aspects of the RNA assessment that can be used as indicators for the quality and purity of the extracted RNA. Ultra violet (UV) spectrophotometry at the following wave lengths can inform on a number of parameters; 240 nm (background absorption and possible contamination), 260 nm (specific for nucleic acids), 280 nm (specific for proteins), and 320 nm (background absorption and possible contamination) (Fleige S and Pfaffl MW, 2006). Sambrook J et al., (1989) suggested that for DNA or RNA quantification, the readings should be taken at wavelengths of 260 nm and 280 nm. The ratio between the reading at 260 nm and 280 nm (A_{260}/A_{280}) provides an estimate of the purity of the nucleic acid in which pure preparation of DNA and RNA have A_{260}/A_{280} values of 1.8 and 2.0, respectively (Sambrook J et al., 1989). A review by Fleige S and Pfaffl MW, (2006) suggested that pure RNA should be 1) free of protein (absorbance 260 nm/280 nm), 2) free of genomic DNA, 3) non-degraded (28S:18S ratio should be roughly between 1.8 and 2.0, with low proportion of short fragments), 4) free of enzymatic inhibitors, 5) free of any substances which complex essential reaction co-factors, like Mg^{2+} or Mn^{2+}, and 6) free of nucleases for extended storage. Pure RNA will have A_{260}/A_{230} equal to A_{260}/A_{280} and >1.8 (Sambrook J et al., 1989).

A new tool for RNA quality assessment is the RNA Integrity Number (RIN) generated by the Bioanalyzer 2100 (Mueller O et al., 2004) (Agilent Technologies). This Lab-on-a-Chip capillary gel-electrophoresis system separates and detects RNA samples via laser induced fluorescence detection. The bioanalyzer software generates an electropherogram which provides a detailed visual assessment of the quality of an RNA sample (Mueller O et al., 2004). Integrity of the RNA may be assessed by visualization of the 18S and 28S ribosomal RNA bands (Fleige S and Pfaffl MW, 2006, Mueller O et al., 2004). The RIN was developed to remove individual interpretation in RNA quality control (Mueller O et al., 2004). It takes the entire electrophoretic trace into account. The RIN software algorithm allows for the classification of riboeukaryotic total RNA, based on a numbering system from 1 to 10, with 1 being the most degraded profile and 10 being the most intact (Mueller O et al., 2004).

Although use of the Bioanalyzer has become accepted in the standardization of RNA integrity assessment, the interpretation of the electropherograms requires some expertise (Imbeaud S et al., 2005, Raman T et al., 2009). Therefore, a common practice is to set up a standard baseline in order for consistent application of the RIN score to ensure reproducibility and reliability of the data. Fleige S and Pfaffl MW., (2006) suggested that a RIN higher than five indicates a good total RNA quality whilst the RIN higher than eight indicates a perfect total RNA for downstream application (Fleige S and Pfaffl MW, 2006). Indeed, a minimum cut-off for RIN score of $\geq$7.0 has been adopted in standard operating procedures at Ark Genomics (Roslin Institute, UK) to ensure reproducibility and reliability of the microarray data for downstream application (R Talbot, personal communication).

In a study by Raman T et al., (2009) samples with RIN $\geq$ 7.0 generally performed well (failure rate of 2.8%) in microarray assessment of gene expression in human airway epithelium and a consistent application of RIN $\geq$ 7.0 as a cut-off was useful for

obtaining high quality gene expression microarray data (Raman T et al., 2009). The effect of sample type on RIN was examined in a variety of calf tissues, white blood cells and four cell lines (Fleige S et al., 2006). For solid tissues the average RIN was between 6 and 8. Tissues or organs with high connective tissue content, e.g. rumen, omasum and jejunum in the gastrointestinal tract, showed increased RIN variation (Fleige S and Pfaffl MW, 2006, Fleige S et al., 2006) likely due to RNA degradation through the sampling and extraction procedures. This may relate to the solid and tough structure of the connective tissues, presence of fatty tissue, the RNase enzymatic activity and/or problems during tissue sampling and storage (Fleige S and Pfaffl MW, 2006). In the context of lung tissue, the average RIN score obtained was 6.55 (Fleige S and Pfaffl MW, 2006, Fleige S et al., 2006).

The RNA samples used in the previous chapter (Chapter 3) came from cartilaginous-tissue samples and little is known about RNA integrity since the quality of the isolated RNA relied only on conventional spectrophotometric absorbance readings (A_{260}/A_{280} and A_{260}/A_{230}). Therefore, this study was designed to determine whether the yield and quality of RNA obtained from these bronchial samples was sufficient for microarray analysis and if not, to investigate whether modifications to our standard extraction protocol could sufficiently enhance the quality of the RNA for this purpose.

A challenging task in microarray-based studies is to extract the valuable information from thousands of genes available in the array. Murphy D, (2002) described two levels involved in microarray data analysis once the primary image data has been collected. The aims of the first level of analysis, so-called low-level analysis, are background elimination, filtration, and normalization, all of which should contribute to the removal of systematic variation between chips, enabling group comparisons. Microarray data normalization involves comparing different microarrays relative to some standard intensity value. Data are often then subjected to log transformation to improve the characteristics of the distribution of the expression values (Murphy D, 2002). High-level microarray analysis, often called "data mining," is the uncovering of relevant patterns of

interest in data and typically, will involve data processing using various statistical techniques to identify the patterns (Murphy D, 2002).

In the present study, the large-scale assessment of gene expression aims to use expression profiling to allow the macro-dissection of molecular events and potentially identify novel pathways and regulators involved in airway regeneration and repair. An improved understanding of the temporal gene expression profiles could be important in understanding the primary mechanisms of cellular dedifferentiation, migration, proliferation, re-differentiation, checkpoint control and repair occurring after airway epithelium perturbation. These findings should open new approaches for determining the possible mechanisms of pathology and repair after lung injury.

5.2 OBJECTIVES

The aims of the study presented in this chapter are divided into two components:

A) Optimising RNA isolation from cartilaginous airway samples

   1) To optimize homogenization techniques for isolating mRNA from lung tissue samples
   2) To identify the amount of the lung tissue required to generate sufficient RNA of the required quality and integrity for microarray analysis

B) Transcriptional profiling of airway epithelial regeneration and repair using microarray

   1) To use single-colour ovine microarray technology to explore the alterations in the transcriptional programme of the airway in this model of airway regeneration and repair.
   2) To identify genes and/or pathways associated with the different stages of the repair process; migration, proliferation and re-differentiation.

5.3 MATERIALS AND METHODS

5.3.1 Optimising RNA isolation

5.3.1.1 Experimental design

A total of 28 tissue samples were obtained from six sheep. These consisted of three different types of lung tissue – cartilaginous airways (n=20), lung parenchyma (n=4) and bronchial brushings (n=4). The amount of cartilaginous airway or parenchyma tissue used as starting material was either 30mg or 100mg. The cartilaginous airway tissues samples were lysed by one of two methods, either by addition of lysis buffer alone (RLT-Buffer; Qiagen, UK) for 30 minutes, or in combination with a homogenisation step using a FastPrep System (MP Biomedicals, UK) and FastRNA ProGreen matrix tube. In the latter regard four different speed/time settings were employed: S_4T_{20} - Speed Setting 4 for 20 seconds, S_6T_{20} Speed Setting 6 for 20 seconds, S_4T_{40} Speed Setting 4 for 40 seconds and S_6T_{40} Speed Setting 6 for 40 seconds. 30mg starting material was used for optimising the FastPrep protocols at different speed (S) and time (T) settings ($S_4T_{20,}$ $S_6T_{20,}$ $S_4T_{40,}$ and S_6T_{40}). For the parenchyma tissue samples, homogenisation followed the protocol suggested by the manufacturer (S_6T_{40}) whilst no homogenisation was required for bronchial brushing samples. RNA was extracted from lysates using RNeasy Mini kits (Qiagen, UK).

5.3.1.2 Isolation of mRNA

Two preservation methods were used: 1) RNAlater: Cartilaginous airway (n=2) and parenchyma (n=4), and 2) Snap frozen: Cartilaginous airway (n=18) and bronchial brushing (n=4). Samples (cartilaginous airways and lung parenchyma) were weighed into a FastRNA ProGreen matrix tube containing 0.7 ml RLT buffer and β-mercaptoethanol (1ml RLT buffer: 10µl β-mercaptoethanol). Samples from bronchial brushings were extracted directly from the cell pellet with RNeasy kit (Qiagen, UK). The tissue samples that were treated only in lysis buffer alone were incubated for 30 minutes in RLT Buffer (Qiagen, UK) and mixed by inverting every 10 minutes.

The upper phase from each FastPrep tube was transferred to a Qiashredder column (Qiagen, UK). For the sample treated only with lysis buffer, the whole 0.7ml buffer was transferred into the Qiashredder column. Samples were centrifuged at 13,000 rpm for 2 minutes and the supernatant transferred into RNeasy columns (Qiagen, UK). The extraction protocol was completed as suggested by the manufacturer. The RNA was eluted using 40µl of RNAse free water (Qiagen, UK).

5.3.1.3 RNA quality check

Measures of RNA quality (purity and integrity) and abundance were determined for each sample using a NanoDrop Spectrophotometer (Thermo Fisher Scientific Inc, UK) to determine the A_{260}/A_{280} ratio and RNA concentration, and the Agilent Bioanalyzer 2100 (Agilent Technologies, UK) to generate an electropherogram and an RNA Integrity Number (RIN).

5.3.1.4 Statistical analysis

To assess the correlations between RIN score, RNA concentration and OD readings, the data for all samples (n=28) were pooled. The mean and standard deviations of RIN score, RNA concentration and A_{260}/A_{280} reading for each tissue type (parenchyma, bronchial brushing and cartilaginous airways) were calculated. The degree of association between the RNA concentration, the A_{260}/A_{280} ratio and the RNA integrity number obtained by the Agilent Bioanalyzer 2100® were analyzed by Pearson correlation coefficient test. A $p<0.05$ was considered statistically significant.

5.3.2 Microarray experiment

5.3.2.1 Experimental design

The design of the experiment carried out for microarray analysis was described in the (Chapter 4; Section 4.3.1).

5.3.2.2 Tissue samples and processing

At post mortem, the lung segments were carefully mapped and dissected out to identify the injured sites, and each injured airway site was immediately cut into 4 or 5 pieces, frozen in the dry-ice/isopropanol and stored at -80^{o}C. Later the tissues were transferred to the RNAlater and kept at -80^{o}C.

5.3.2.3 Total RNA isolation

Total RNA was isolated as described above but using 100mg of cartilaginous airway tissue a FastPrep setting of $T_{40}S_6$. The rest of the RNA isolation procedure was performed using RNeasy Mini kits (Qiagen, UK) following the manufacturer's instructions. The RNA was eluted using 40µl of RNAse free water (Qiagen, UK).

5.3.2.4 RNA quality control assessment

The quality and yield of the extracted RNA was initially measured with a NanoDrop ND-100 Spectrophotometer (Thermo Fisher Scientific Inc, UK) to determine the RNA concentration and A_{260}/A_{280} ratio. Samples were subsequently sent to the Ark Genomics (Roslin Institute, UK) for analysis using an Agilent Bioanalyzer 2100 to generate electropherogram profiles and to assess the RIN score.

All the procedures used in this protocol were followed the techniques as recommended by the manufacturer. The Agilent 2100 Bioanalyzer with the RNA 6000 LabChip kit (Agilent, UK) was used as a more stringent assessment of mRNA quantity, integrity and purity. This system uses fluorescence detection, monitoring the fluorescence between the wavelength of 670 and 700nm for RNA. RNA samples are normally considered good quality and integrity if the RIN produced is greater than 7.0 and two distinct peaks for 18S and 28S ribosomal RNA are clearly observed.

5.3.2.5 Designing of oligonucletide probes

In our preliminary study (Chapter 3) a semi-quantitative RT-PCR approach was used to look at expression of selected genes in a limited number of sheep. The microarray study

gave us the possibility of looking at the same genes (except keratin 14 gene) in a more quantitative manner in a larger number of animals. However, of these genes, the array only contained a probe for the beta-catenin gene; catenin beta-like 1 (CTNNBL1); Probe: A_70_P068021 (Accession No: DQ152928). Additional probes were designed by Ark Genomics, Roslin Institute UK. To design probes for beta-tubulin, keratin 5, and FoxJ1 the sequences generated in Chapter 3 were used to Blast through the GenBank and DFCI databases to find similarities with ovine and/or bovine sequences. Since there was no available ovine sequence for FoxJ1 gene, both the ovine sequence generated in Chapter 3 for FoxJ1 gene and the bovine sequence for FoxJ1 gene (btFoxJ1_XM582255.f1) were used – (Table 5.1)

Probes Name	Probes Sequence
Keratin5_TC8502.f1	CCCTCTCCTCCAATTACTTCCTCTCTTCTGCTCCCTTCTTCTCTTACTCCTT
btFoxJ1_XM582255.f1	CTTCCTGGCCACGTCCTTCCTGCAGCACCCTTGGGATGAGAGCAGCAGTA
TubulinB_TC2191.f1	CAATCCCCTCAGCCTTCTTCCTCAACGACCCCTTTCCTCTCCCTCAGAAT
Sheep_Foxj1.f1	GATGTTGACCGCCCCCTCGGAGAGTCCCCGCAGATATGGCGGAGAGCTGG

Table 5.1: Lists of probe sequences added to the array
These probes were designed based on similarity between sequences generated in the previous preliminary study (Chapter 3) and bovine and/or ovine sequences available on the DFCI website. For beta tubulin, when the probes were designed, blast analysis of the sequence generated in the Chapter 3 was similar to this EST sequence (TC2191). Similarly, there was no ovine FoxJ1 mRNA sequence available when the probes were designed; therefore both FoxJ1 sequences generated in the Chapter 3 and bovine FoxJ1 mRNA sequence (btFoxJ1_XM582255) were included in the array.

5.3.2.6 RNA Amplification, cDNA synthesis, labelling, hybridization and scanning

For the microarray experiment, all the procedures including RNA amplification, cDNA synthesis, labelling and hybridization, were carried out by the professional technician at Ark Genomics, Roslin Institute, UK following Ark Genomics Standard Operation Procedures (SOP) – please see Supplement 7a for the details (on the CD).

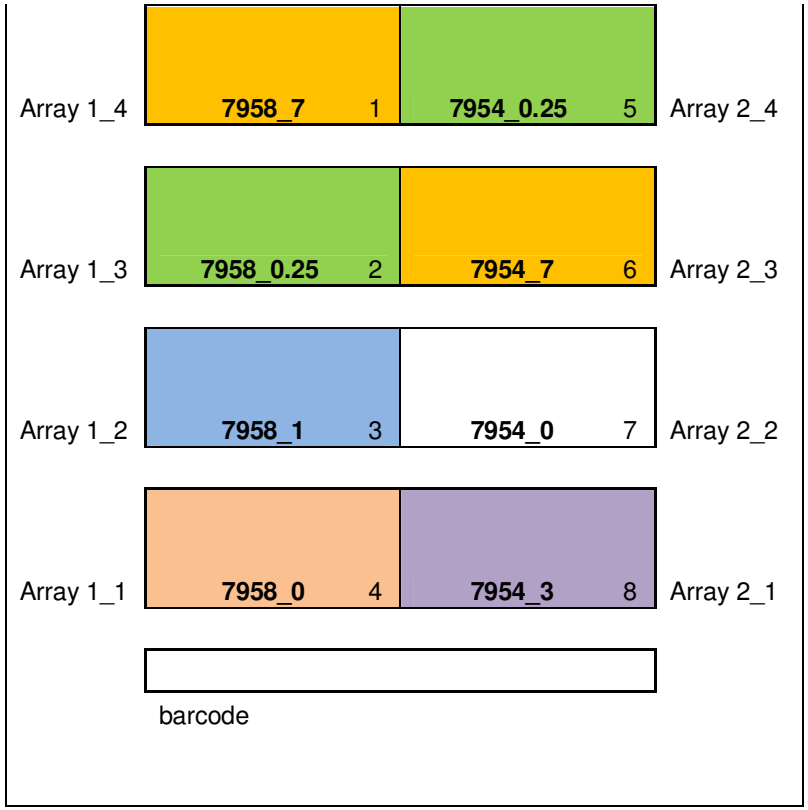
Four slides containing eight arrays (n=8) each were used for the microarray study. As described previously, there were six replicates (n=6) for each of the four time points which were all taken from different sheep (Table 4.1). One segment from each of the treated sheep was used as a naive control (n=8 naive controls). Sheep were allocated at random to Agilent slides and samples were allocated at random to arrays within slides so that any variation between slides or arrays caused by the measurement process should not bias the comparison of treatments. The allocated positions of each of the samples within the slides are shown in Figure 5.1.

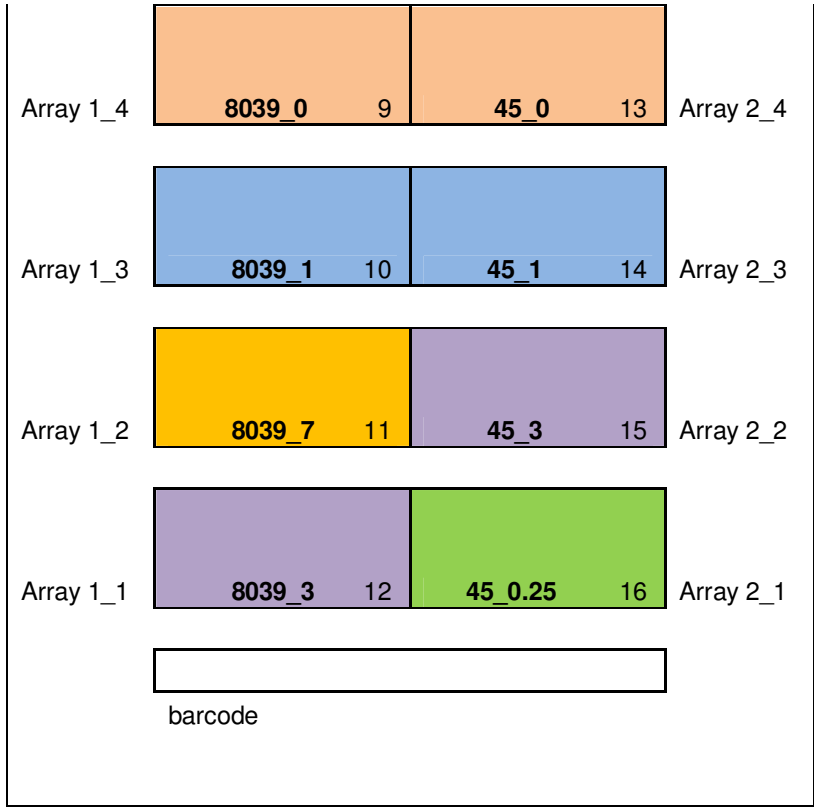
The glass slides containing the RNA samples were scanned using the auto photo multiplier tube (PMT) Axon 4200AL scanner following the manufacturer's standard operation procedures. The microarray expression data was extracted using the Agilent Feature Extraction (FE) software version 9.5.3.

When the Agilent microarray file generated by the scanner is loaded into a project, the software assigns a grid template based on the bar code (AMADID), and a protocol. The protocol uses the grid template to automatically locate the spots. After the software positions the grid and finds spots, it uses the remaining default FE algorithm parameters contained in the FE protocol to carry out data analysis. See supplement 7b on the CD for the FE parameters table. Feature Extraction removes outlier pixels and does statistics on inlier pixels of features and local backgrounds (FE Software Manual, Agilent version 9.5.3 for one colour). The gMeanSignal used in this anlysis is one of the basic values generated by the FE software. Raw signal is calculated, based only on the "inliers" pixels, meaning all the pixels in a spot retained

after removing saturated, or not "well above background" pixels. The CookieCutter method was used to select all the pixels used in the calculation.

Quality control checks were performed on the microarray data output generated from FE software. The quality control parameters followed were as recommended by the manual for the Agilent Feature Extraction Software (version 9.5.3) for one colour.


Array 1_4
7958_7 1
7954_0.25 5
Array 2_4
Array 1_3
7958_0.25 2
7954_7 6
Array 2_3
Array 1_2
7958_1 3
7954_0 7
Array 2_2
Array 1_1
7958_0 4
7954_3 8
Array 2_1
barcode


Array 1_4
8039_0 9
45_0 13
Array 2_4
Array 1_3
8039_1 10
45_1 14
Array 2_3
Array 1_2
8039_7 11
45_3 15
Array 2_2
Array 1_1
8039_3 12
45_0.25 16
Array 2_1
barcode

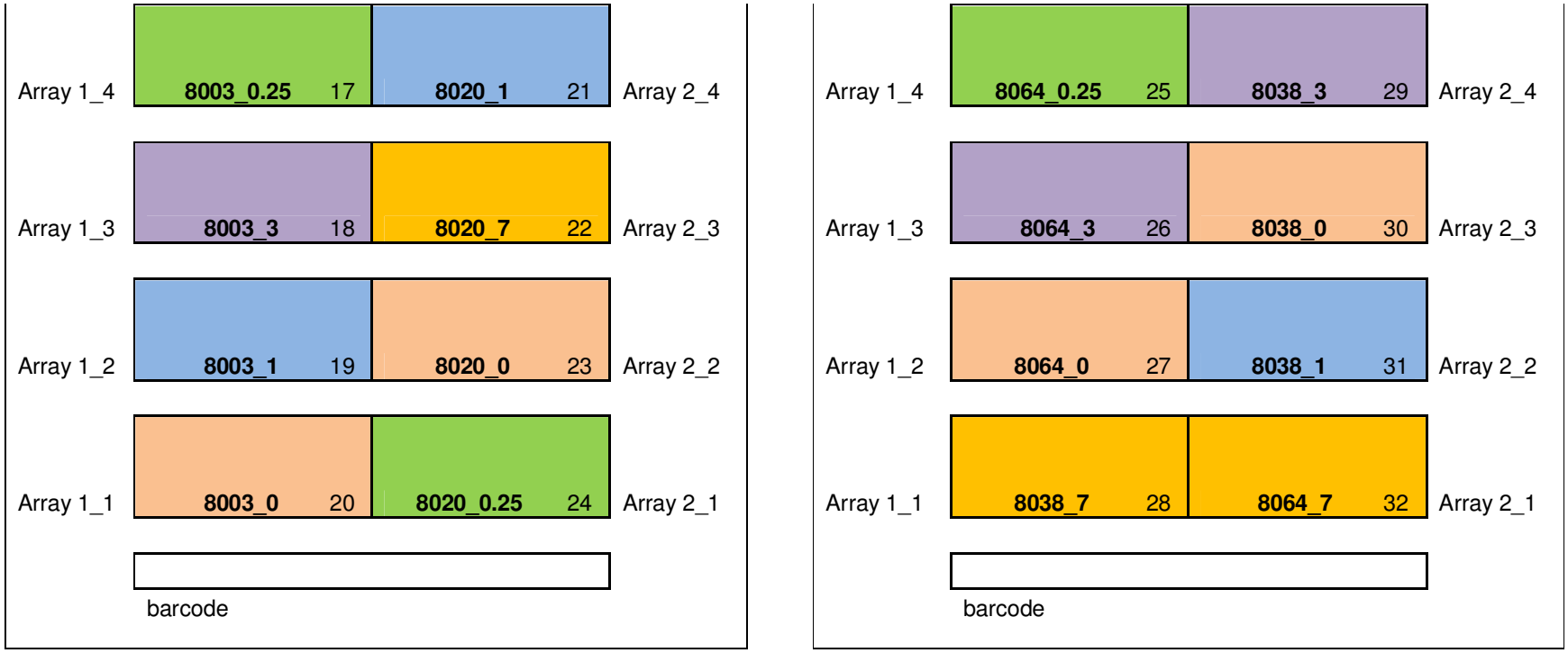


Figure 5.1: The allocated array positions for the samples on each of the slides.
Sheep were allocated at random to Agilent slides and samples were allocated at random to arrays within slides. The numbers in each box indicate sheep ID, time point and, on the right hand side, the numbering system (slides position) given to the arrays.

5.3.2.7 Microarray statistical analysis

The study investigated 5 different time points following injury (Control (uninjured), 6 hours, 1 day, 3 days and 7 days as described in Chapter 4; Section 4.3.1.) from four combinations of time points (protocols) at which each sheep was injured (Table 4.1). Therefore, in order to assess the effect of the brushing on the gene expression profiles over the time of injury, the effects were estimated by comparing each of the four time points between control (naïve) samples. To compensate for an interaction between the time of the brushing (Time point) and the protocol (combination of time points), brushed sites were compared to naive control airway from animals subjected to the same protocol. For instance, for protocol 731, there were two sheep which were subjected to brushing at 7, 3 and 1 days prior to harvest. In order to assess the effect of the brushing on the gene expression profiles, the data generated from the microarray from each time point (day 1, day 3 and day 7) were only compared with the microarray data generated from control samples of the same two sheep (n=2)

Analysis of microarray data was performed using a linear mixed model fitted by residual maximum likelihood (REML) (Patterson and Thompson, 1971). For this analysis, the gMeanSignal (raw means signal of feature in green channel) from microarray raw data extracted from Agilent Feature Extraction software was used and gMeanSignal intensity was transformed to logarithms base 2 (logBase2) for each probe separately using GenStat commands and procedures (Payne et al., 2008).

5.3.2.8 Microarray validation study

Semi-quantitative RT-PCR was performed on selected genes from the microarray experiment in order to validate the observed changes. Four samples for each time point (total n=20 samples) were selected and used in this validation. The selection of these genes was based on the gMeanSignal values with less variation whilst values with large variation were excluded. In addition, the genes selected for the assay were based on 1) a significant change in expression (up or down) at one or more time points at $p<0.001$ (highly significantly changed), 2) either high or low fold change at one or more time points, 3) genes thought to be have a relevant biological function such in cell cycle, immune response, cell communication, cell adhesion and cellular

differentiation, and 4) genes described in Chapter 3. The full list of genes used for RT-PCR assay is listed in Table 5.2. The semi-quantitative RT-PCR protocols used were similar to those described in Chapter 3. Data for these genes were normalised to ATPase gene expression.

The gene annotations for the probes on the sheep array were obtained by blasting the mRNA sequences of these genes through the nucleotide mRNA database using the NCBI programme available at http://blast.ncbi.nlm.nih.gov/Blast.cgi. The selection criteria for selecting the annotations for the sequences of interest were based on a sequence similarity of more than 90% with bovine mRNA sequence database. The mRNA sequences for these genes were used for primers design using http://biotools.umassmed.edu/bioapps/primer3_www.cgi. The list of the primers and the PCR conditions are shown in (Table 5.2).

In order to assess the correlation between expression data obtained from the array and semi-quantitative RT-PCR, a Pearson correlation coefficient test was carried out using R program with $p<0.05$ considered significant. The array data for some of the genes were log-transformed due to the fact that they were not normally distributed. Expression profiles of the array and semi-quantitative data of the selected samples were plotted along with the scatter plots for the correlation test.

Genes	Primers	Sizes (bp)	Tm (^{o}C)
CCNB2	Forw: CTGTGTCCAGTCCGAGTCCT Rev: GCACTGGTTTCACAGAAGCA	304	59
CCNB1	Forw: ACCTCTGGAAAAGGCTCCTG Rev: GGTTGACTGCTTGCTCTTCC	303	59
CDC2	Forw: CTCCACAGGGACCTAAAACC Rev: AGTGCCCAAAGCTCTGAAAA	294	59
KRT6A	Forw: TACACCACCACCTCCTCCTC Rev: GCCAGTGGAAAGCAAGTAGG	299	60
KRT19	Forw: CTGGAGAACAGCCTGGAAGA Rev: GCACGTACACGACTTTGGTG	300	60

Table 5.2: The list of the genes for a microarray validation study using a semi-quantitative RT-PCR assay.
(CCNB2; cyclin B2, CCNB1; cyclin B1, CDC2; cell division cycle 2, KRT6A; keratin 6A, KRT19; keratin 19)

5.4 RESULTS

5.4.1 Optimizing RNA quality and integrity

5.4.1.1 Preliminary observations

Our preliminary data for quality control assessment of the RNA samples for microarray analysis indicated that of the 12 samples screened for RNA integrity using Agilent Bioanalyzer 2100, only 1 had passed the RIN cut-off (≥7.0). Although similar techniques for RNA isolation and QC assessment were repeated several times using the same and/or different samples all with high RNA concentration, there was no improvement observed in terms of the RIN score. To preserve the remainder of these samples for the microarray study, a set of normal cartilaginous airways tissues were obtained to optimize the RNA extraction procedures.

5.4.1.2 RNA quality from airway samples

Initial RNA was extracted from 30mg of cartilaginous airway samples as recommended in the RNeasy Mini Kit Protocol. For these samples, the average ratios for A_{260}/A_{280} were ≥1.8 – (Table 5.3). However, the RNA concentrations obtained from these tissues were considered suboptimal in the sense that the Agilent Bioanalyzer (Mueller O et al., 2004) specifies a minimum RNA yield of 25ng/µl for

a quality control (QC) check and approximately 10µg - 20µg of RNA is required for microarray (Wang S et al., 2003). From 30mg tissue, the average RIN score obtained was very low (RIN <4.0) as compared to the cut-off (≥7.0). The electropherograms (Figure 5.2; left panel) did not have the desired flat baseline with distinct 18S and 28S ribosomal RNA peaks expected from good quality RNA (Figure 5.2; right panel). Figure 5.3 shows the different of 18S and 28S ribosomal RNA peaks profiles from the various RNA integrity (RIN) scores screened on the cartilaginous airway tissue samples.

Increasing the amount of starting material to 100mg improved the RNA yield. The average ratios for A_{260}/A_{280} were ≥ 1.8 – (Table 5.3), even though the amount of the tissue exceeds that recommended by the RNeasy Mini Kit Protocol. The average RIN score also increased (RIN = 6.0) but this was still lower than target (RIN > 7).

	30mg (n=7)		**100mg (n=7)**	
	Mean	**±SD**	**Mean**	**±SD**
Yield (ng/µl)	7.93	6.65	57.61	1.41
$A_{260}A_{280}$ (nm)	2.13	0.64	2.02	0.07
RIN	2.77	3.08	6.0	1.41

Table 5.3: Mean values for samples types and the amount of tissues used
There was an effect of the amount of starting material for isolating the RNA from cartilaginous airway tissues. The RNA quantity and integrity was improved with 100mg tissue.

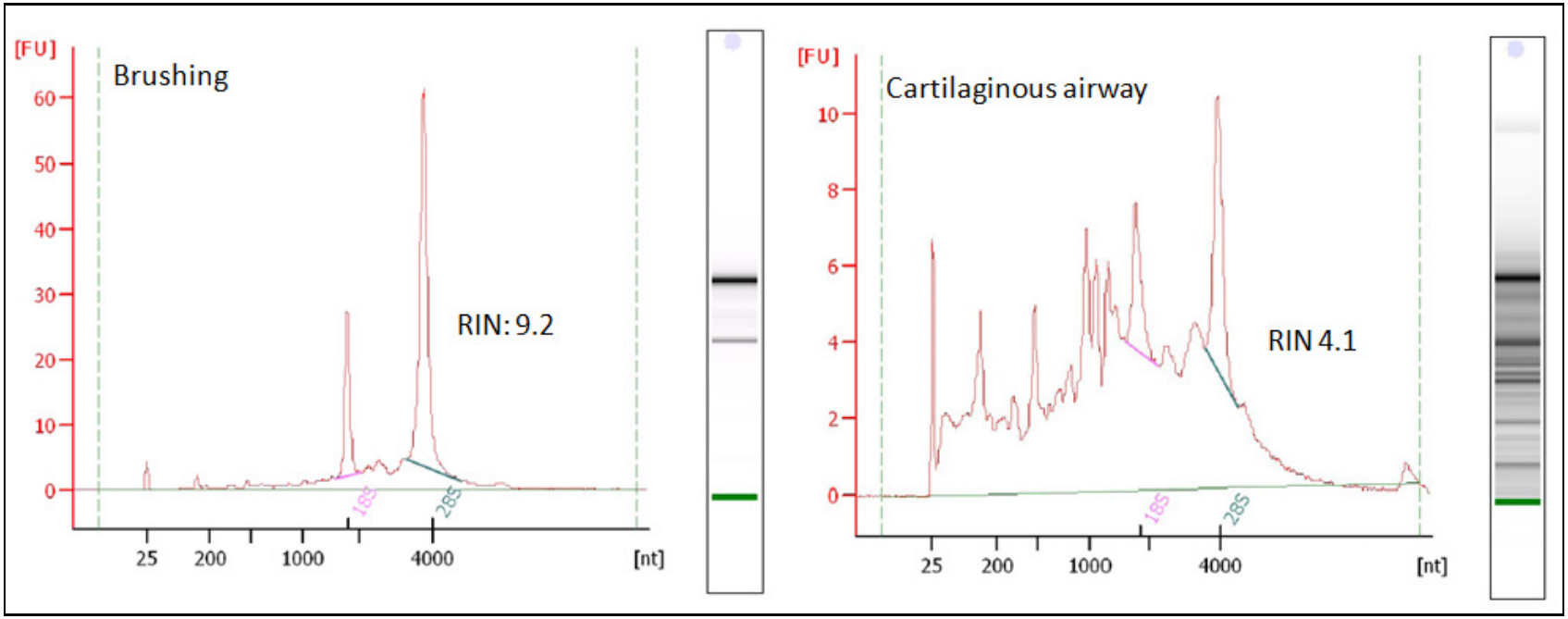


Figure 5.2: Electropherograms demonstrating profiles associated with both low and high RNA integrity number.
Quality and integrity of RNA assessed by the Agilent 2100 Bioanalyzer. For the sample with high RIN a flat baseline and distinct peaks of both 18S and 28S of ribosomal RNA are clearly seen (left panel) while the sample that had low RIN score (right panel) has a noisy baseline that appeared early before the appearance of the 18S rRNA and is indicative of degraded RNA. In the cartilaginous RNA sample, the RNA concentration is low (Y-axis) and makes the background degradation appear high. The X-axis represents the size of fragments (nt: nucleotides) and the Y-axis represents the wavelength of the fluorescence (FU: fluorescence unit).

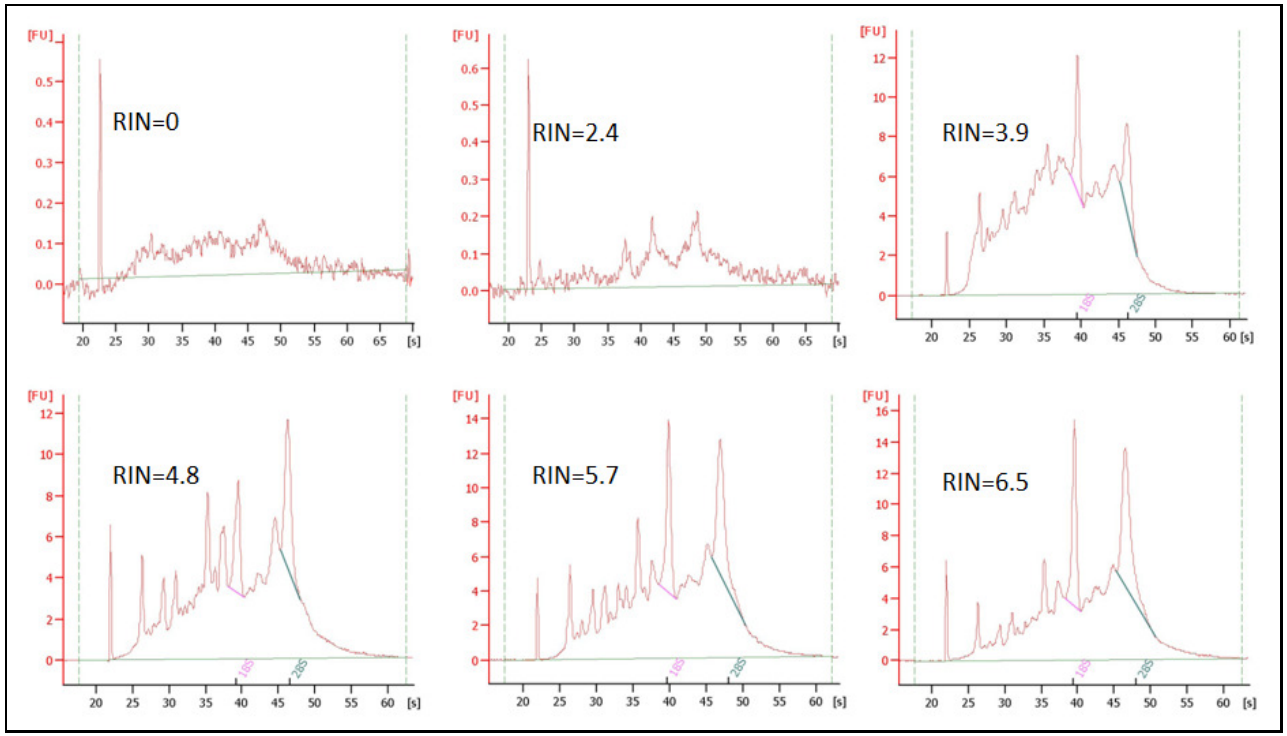


Figure 5.3: The differences of 18S and 28S ribosomal RNA peaks profiles from the various RNA integrity (RIN) scores screened by the Agilent Bioanalyzer 2100
Although additional peaks reflecting RNA degradation can be appreciated clear 18S and 28S ribosomal peaks could still be discerned in these samples. The gradual degradation of rRNA is reflected by a continuous shift towards shorter fragment sizes.

5.4.1.3 Homogenisation parameters

Variations in the speed and time settings for the FastPrep system were compared to address the concern that our standard settings were resulting in RNA degradation. Within the 30mg samples of cartilaginous airway tissue samples assessed, there was no improvement in the RIN score achieved for the samples prepared at any of the settings tested (Table 5.4).

Samples	FastPrep setting	Tissue amount (mg)	RNA conc. (units)	A_{260}/A_{280}	RIN
Cartilaginous airways	lysis buffer	30	6.05	2.01	0
Cartilaginous airways	lysis buffer	30	8.42	2.03	0
Cartilaginous airways	S4/T20	30	1.79	2.65	0
Cartilaginous airways	S4/T20	30	2.17	2.06	2.4
Cartilaginous airways	S4/T40	30	9.73	1.69	2.5
Cartilaginous airways	S4/T40	30	5.4	2	2.5
Cartilaginous airways	S6/T20	30	8.78	1.79	2.3
Cartilaginous airways	S6/T20	30	4.96	2.05	2.4

Table 5.4: The effects of different processing protocols on the quality and integrity of the RNA from the cartilaginous airway samples
The results indicate that the small amount of the initial tissue samples used in RNA extraction rather than the Fastprep conditions had contributed to the low RIN score. Different homogenising settings for the cartilaginous airway samples failed to make an impact on the overall yield of the extracted RNA (possibly due to low cellular density in this tissue type).

5.4.1.4 Lung specific or sample type specific?

To test the hypothesis that this problem was specific to the cartilaginous airway samples, RNA was extracted from other lung-derived samples, namely bronchial brushings and lung parenchymal tissue. The A_{260}/A_{280} ratios from both these sample types were consistent and generally equivalent and indicated acceptable RNA purity. RNA from these samples generated improved RIN values (Table 5.5), indeed all of the bronchial brushing samples demonstrated a RIN score greater than 7 with clear

and distinct peaks for both 18S and 28S ribosomal RNA (exemplified by the right panel in Figure 5.2), despite the low RNA yield.

Tissue types	Average(±SD)		
	RNA conc.(ng/µl)	A_{260}/A_{280}	RIN
Parenchyma (n=4)	114.34 ±89.45	2.08 ±0.05	7.34 ±0.79
Bronchial brushing (n=4)	18.23±4.69	2.09±0.09	8.55±0.79

Table 5.5: A summary of the RNA concentration, A_{260}/A_{280} ratios and the RIN scores for the parenchyma and bronchial brushing samples
Mean (+/- SD) values for parenchyma and bronchial bushing samples. RNA concentration, A_{260}/A_{280} ratios and, the RIN score obtained

5.4.1.5 Is RNA integrity dependent on RNA yield and purity?

To assess the individual tissue type effect on RNA concentration and RIN score, data was plotted based on the initial tissue weight (30mg vs 100mg for both parenchyma and cartilaginous airway tissue types), RNA concentration and RIN scores (Figure 5.4). This figure shows that although the RNA concentration isolated from bronchial brushing and parenchyma tissue samples were low, the majority of RIN scores for these samples were greater than 7. In order to assess a correlation between RNA yield and purity of the samples, the data of each variable (yield and purity) for all samples were combined. There was a positive correlation (Figure 5.5; left panel) observed between the RNA concentration and the RIN score for the same samples (Spearman r_s=0.64, p<0.001). There was no correlation observed (Figure 5.5; right panel) between the A_{260}/A_{280} ratio and the RIN score for the same samples (Spearman, r_s=0.23, p>0.05)

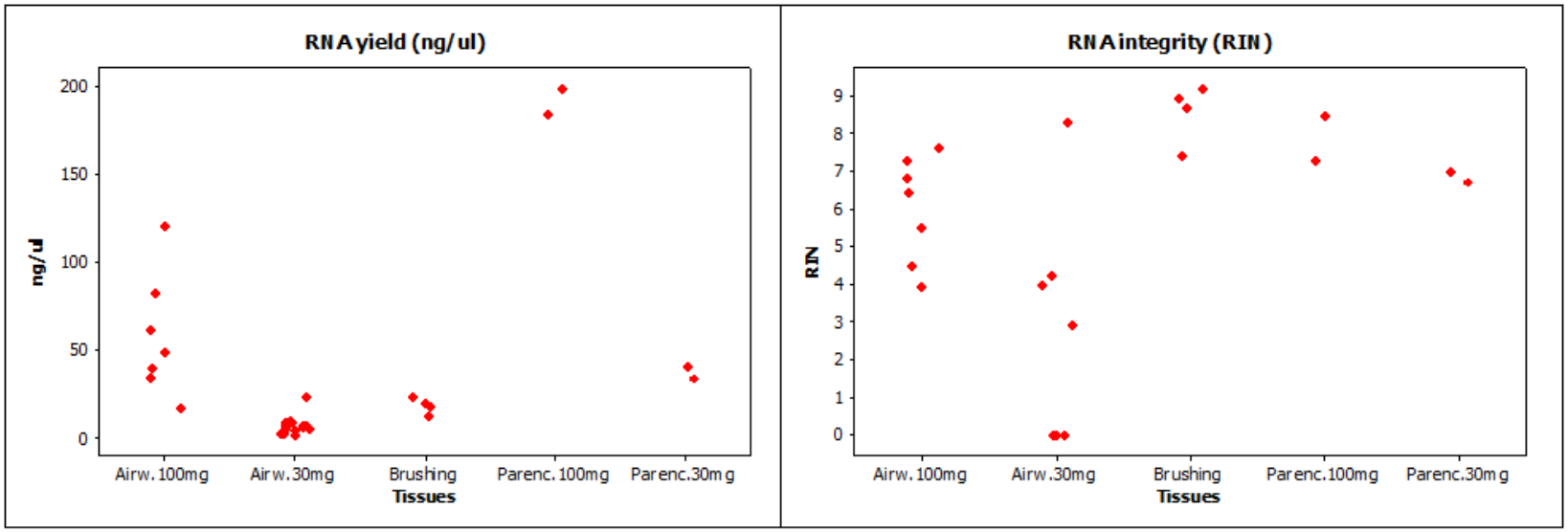


Figure 5.4: The effect of different tissue types on the RNA quality and integrity.
There were low RNA concentrations obtained from the 30mg samples of all lung tissue types. Despite the low RNA concentrations obtained for the parenchyma (30mg) and bronchial brushing, these samples reached higher RIN values than the RNA from 100mg of cartilaginous airway tissue.

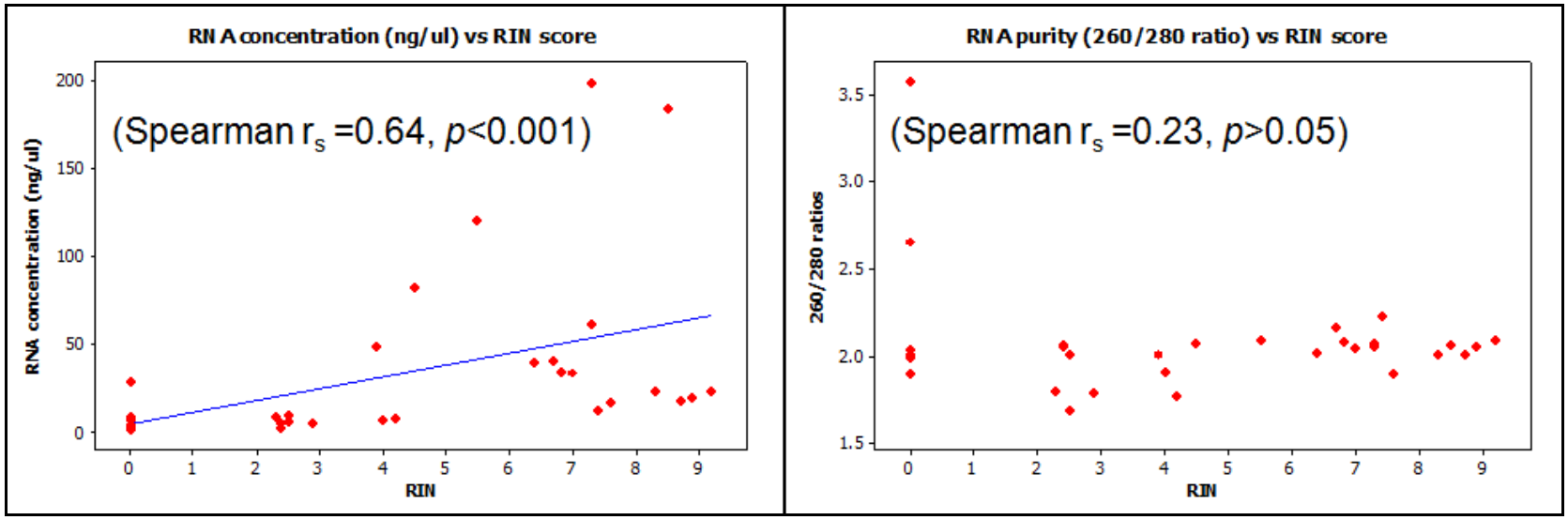


Figure 5.5: The relationship between RNA concentration and A_{260}/A_{280} ratio.
The scatter plots show the relationship between RNA concentration (left panel) and A_{260}/A_{280} ratio (right panel) with RNA integrity number obtained. The data of each variable (yield and purity) for all samples (n=28) were combined. Overall, there was evidence of a significant association between the yield of the RNA and the RIN number (Spearman r_s=0.64, p<0.001) whilst no such association could be demonstrated between the A_{260}/A_{280} ratios and RIN score (Spearman, r_s=0.23, p>0.05).

5.4.2 Analysis of gene expression study

5.4.2.1 RNA quality control check

In total, there were 32 samples were subjected to microarray experiment consists of 8 naive controls and 24 treated samples that were subjected to brushing at various time points. The details of the samples and their RIN scores, initial RNA concentration, aRNA volumes for hybridization process are shown in the Supplement 8 on CD. The first quality control check for the RNA samples was carried out immediately after isolation using the conventional spectrophotometric method of RNA quantification. The A_{260}/A_{280} ratios for all samples were within the "very good" range of RNA purity between of 1.9 to 2.15 while the average RNA concentration was 109.2±77.67 ng/µl. The second quality control check was performed using the Agilent 2100 Bioanalyzer to assess the RNA integrity and quality. An average RIN score of 7.42 ± 0.63 was achieved. For a number of samples, the Agilent recommended threshold for microarray analysis (RIN>7.0, Talbot R, personal communication) could not be achieved, even after several optimisation steps, therefore the minimum RIN score accepted for subsequent microarray analysis was lowered to 6.0. The appearances of the 18S and 28S ribosomal peaks were also taken into consideration. During the quality control check, overall, 4/6 samples belonging to the day 1, 2/6 from day 3 and 1/6 samples each for days 0 and 7 post injury groups had a RIN score < 7. Two of these had RIN score <6.5 (RIN 6.0 for 6-hour sample and RIN 6.2 for day 1 sample). Overall 94% (n=30) of the samples had a RIN score more than 6.5. A Pearson correlation coefficient test demonstrated that neither RNA concentration nor A_{260}/A_{280} ratios were predictive of the RIN for the RNA samples (p>0.05) – (Figure 5.6). Similarly, there was no correlation (p>0.05) between the time from injury and the RIN number obtained for RNA samples – data not shown.

The third quality control check was at the amplification stage where the sizes of the amplified RNAs (aRNA) were measured. The required size distribution range for reproducibility and optimal microarray performance is between 300 to 4000 nucleotides (Talbot R, personal communication). At this stage, eight of the amplified samples (Table

5.6) had unacceptable size ranges of approximately 500-1000 nucleotides. This failure did not appear to be dependent on the RNA integrity and RNA concentration (the means of the concentrations of these samples were 49.07±17.81ng/µl). However, increasing both the annealing time and the amount of the *Transcriptase* enzyme for the cDNA synthesis for these eight samples gave improved average sizes of around 2000 nucleotides which are in line with the manufacturer's minimum expectations (Ambion, Applied Biosystems). The average peak sizes for the majority of samples were around 3000 nucleotides.

Samples ID	Time points (days)	RIN
8039-8B	0	7.7
45-3B	1	7.5
8003-8B	1	6.3
8020-4B	0	8.1
8003-3B	0	6.2
8038-8B	0	7.8
7958-8B	1	7.5
8020-8B	0.25	6.3

Table 5.6: Samples which generated oligonucleotides with lower average sizes.
Table shows the RIN for samples which performed poorly in the amplification process and the time points post-injury at which they were harvested.

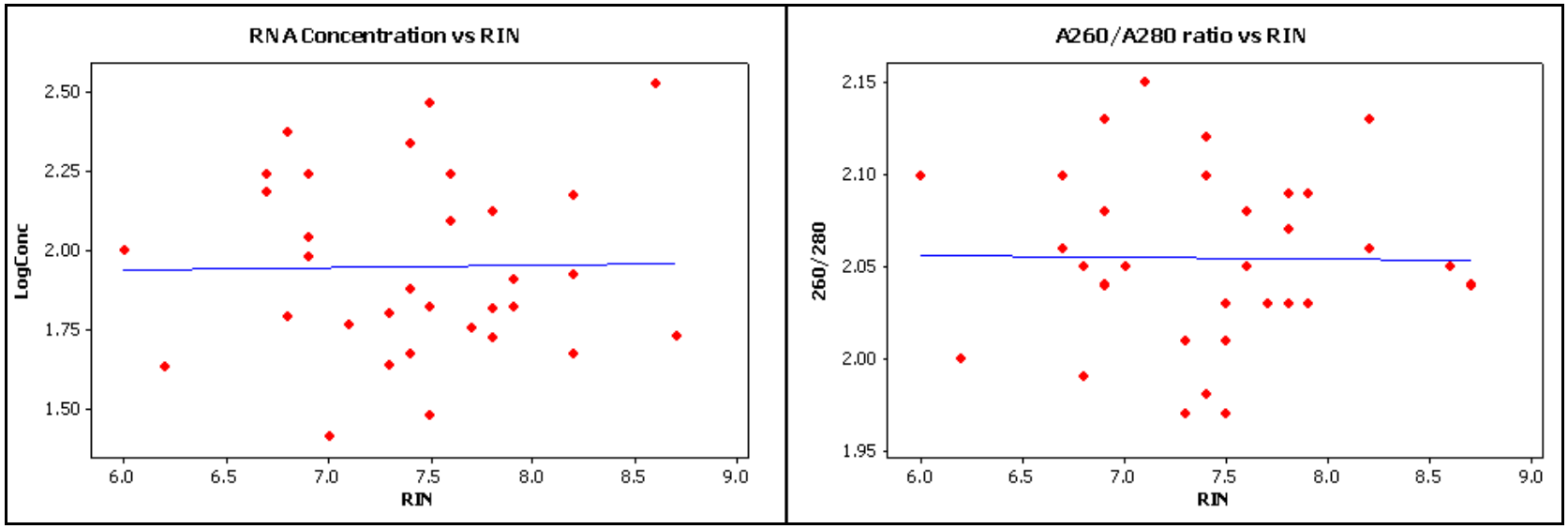


Figure 5.6: An association between the effect of the RNA concentration and RNA purity measured by spectrophotometer with the RNA integrity numbers (RIN) measured by Agilent Bioanalyzer.
Neither both components (RNA concentration –left panel and RNA purity –right panel) have a significant correlation with the RIN score obtained (p>0.05).

5.4.2.2 Quality control of the microarray output

In order to ensure their reproducibility and reliability, the outputs generated from the Agilent Feature Extraction (FE) software were analysed. The QC reports from the FE software output for all 32 arrays are provided in the Supplement 9 on CD. The raw data files for each array (32 arrays in total) are provided in the Supplement 10 on CD. JPEG (Joint Photographic Experts Group standard) images for all 8 slides which contain 32 arrays are shown in the Supplement 9 on CD. The standard QC metrics for one colour arrays include spot finding of the four corners of each array, spatial distribution of the outliers and the histogram of the signals across all non-control probes. The FE output profiles of the arrays used in this study indicated that the microarray data generated were reliable and of good quality. In this regard the quality control reports from the feature extraction showed that there were no spatial abnormalities and the spike in control profiles in both the feature extraction and other analysis programs (i.e GeneSpring) showed that the arrays were performing normally and no array(s) were performing exceptionally in this regard. Indeed the chip spike-in control signal profiles behaved similarly and that there were no obvious outliers. The presence of outliers might have raised questions regarding RNA quality and prompted further analysis to relate RNA quality metrics to array QC metrics. Further data analysis using GeneSpring identified that samples with small fragment sizes tended to cluster separately to the bulk cluster on the principal component analysis (PCA) plot. However QC metrics from those samples were, as previously indicated, considered normal and it was adjudged that those samples did not differ to such an extent that excluding them from further analysis would be warranted (data not shown). Figure 5.7 shows a representative image of the overall distribution of the outliers of the array in which the distribution of the fluorescence signals across the microarray indicated that the hybridization process was achieved in an even and unbiased manner (the rest of the images are shown in Supplement 9 on CD for "distribution of the outliers" and "signals histograms").

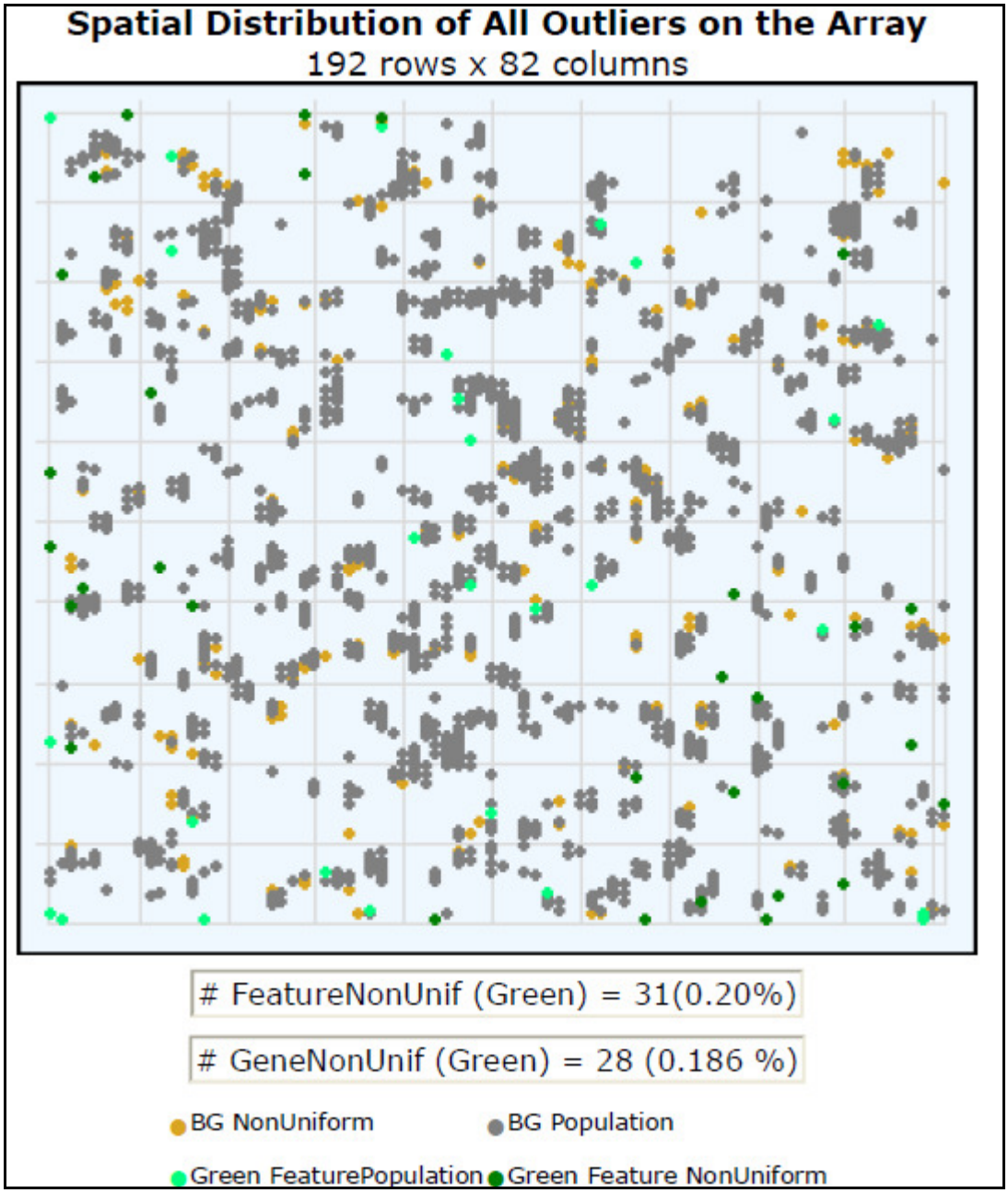


Figure 5.7: A representative image of the spatial distribution of the outliers which extracted from the feature extraction software. This image taken from the slide 1, Array_1_3 and the sheep 7958 (6 hours time point). For one colour platform, only green colour appears. This view is a way to visualize in the spots rejected (background or non uniform) are organized in a certain way. The image shows that the fluorescence signals are distributed across the array.

In this study, examination of the spatial distribution of all outliers on the arrays demonstrated no systematic bias within arrays. In terms of between-array and between-slide normalization, a plot of the log2gMean signals across the four slides, organised according to column and row position on the slide indicates that there is no systematic (ie shared between all slides) bias detectable (Figure 5.8).

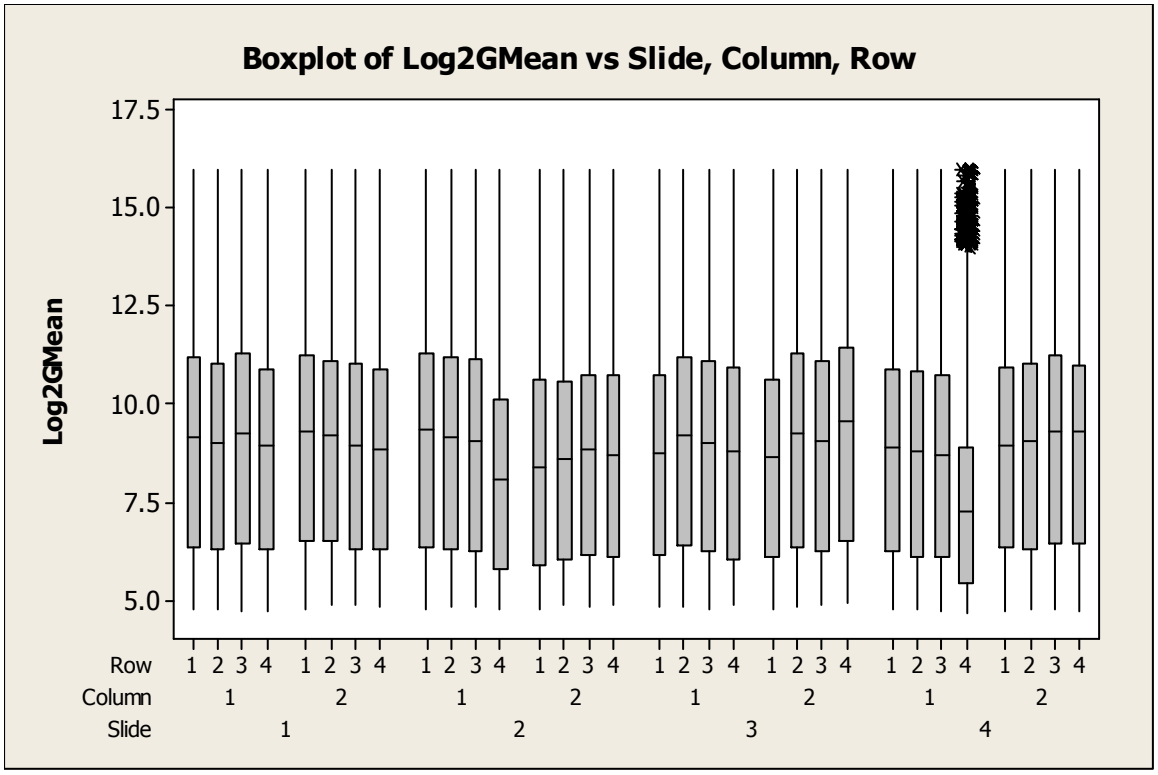


Figure 5.8: A normalization between-array and between-slide. A plot of the log2gMean signals across the four slides, organised according to column and row position on the slide indicates that there is no systematic (ie shared between all slides) bias detectable

The spread of the intensity log ratios is largely similar across slides and importantly there is no evidence of spatial bias. These observations, in conjunction with the additional knowledge that the time points were randomly allocated between rows within columns (sheep), indicate that any detected significant differences in gene expression between time points are more likely to be a true reflection of biological variation and not influenced by systematic procedural bias.

5.4.2.3 Changes in airway gene expression induced by bronchial brushing

The gene expression changes at the brushed sites were evaluated by comparing the gMeanSignal expression values for each of 15015 probes at each of the time points (n=6 samples per time point) with those of the corresponding paired naïve samples (day 0). The expression level of a gene derives from the number of bound mRNA molecules as measured by the green fluorescence intensity in the microarray experiment.

A linear mixed model was fitted by REML to intensity observations transformed to logarithms base 2 for each probe separately using Genstat commands and procedures. In the first analysis, the model included random variation from three sources: between Agilent slides, between sheep within slides and between samples within sheep, and treatment effects, ignoring any interaction between Time points and Protocol. However, there was evidence of an interaction between Time points and Protocol in many probes. The shape of the Time point profile showed variation with Protocol over and above that expected by chance. This is indicated by the backward J-shaped frequency histogram of the significance probabilities (Figure 5.9) associated with the Time point x Protocol interaction for each probe, as described by F statistics derived from Wald statistics (analogous to variance ratios in analysis of variance). Significance probabilities would be expected to follow a uniform distribution (level across the range 0-1) if interaction was zero in all probes.

This interaction could be due to a systematic effect of the timing of injury on the expression profiles over time, which would require further investigation. It could also be a chaotic interaction (i.e. with no coherent pattern) caused by factors involved in the management and treatment of sheep and the collection and processing of samples prior to the microarray process.

Assuming the latter, an additional linear mixed model was fitted which included the Time points x Protocol interaction as a further source of random variation. Estimates of

effect sizes and fold changes remained unaltered by this but standard errors and significance probabilities generally increased in the model with no interaction (Table 5.7). The significant changes observed from the statistical analysis of microarray data are considered up- or down-regulated based on the values of fold changes in which fold change <0.5 (expression down-regulated), >2.0 (expression up-regulated) and 0.5>FC<2.0 (unchanged).

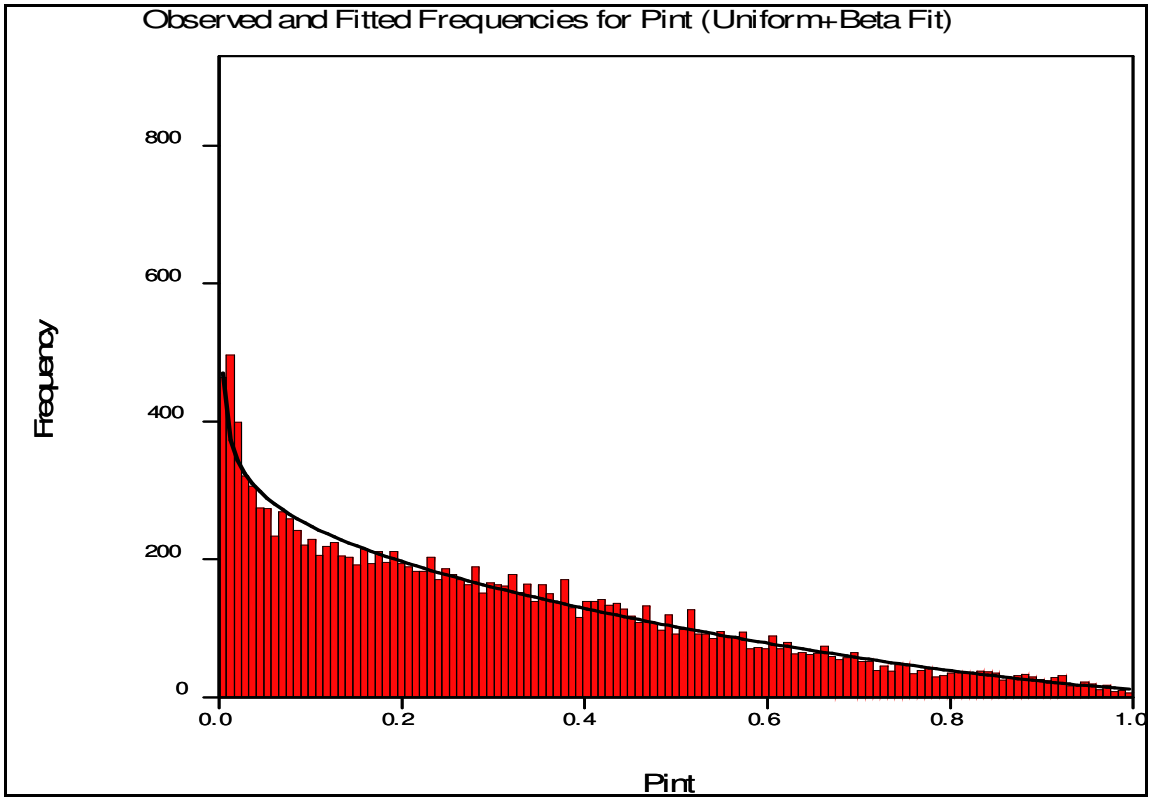


Figure 5.9: Figure shows the backward J-shaped frequency histogram of the significance probabilities of interaction between the time points and the protocols.
Frequency histogram of the significance probabilities associated with the Time point x Protocol interaction for each probe, as described by F statistics derived from Wald statistics (analogous to variance ratios in analysis of variance). Significance probabilities would be expected to follow a uniform distribution (level across the range 0-1) if interaction was zero in all probes. (Pint = *p*-values for interaction)

gMeanSignal : No interaction Time points x Protocol

Time point contrast	% of probes differentially expressed [1]	Number of probes with fold change:		Number of probes achieving statistical significance [2]:		
		>2	<1/2	$p<0.05$	$p<0.01$	$p<0.001$
6 hrs vs Naive	46	301	801	4236	2133	668
1 day vs Naive	27	463	240	2340	1031	395
3 days vs Naive	39	475	413	3040	1345	457
7 days vs Naive	47	292	173	2317	796	206

gMeanSignal : Random interaction Time points x Protocol

6 hrs vs Naive	37	301	801	2659	842	96
1 day vs Naive	15	463	240	1325	435	91
3 days vs Naive	28	475	413	1797	552	77
7 days vs Naive	38	292	173	1175	243	20

Table 5.7: Output of two different statistical models
Two different statistical models were tested in order to compare between probes in treated samples for each time points with naive control samples (time point 0) in which models were fitted with first, with no interaction between Time points and second, with a random interaction. There was a lower percentage of probes differentially expressed when model was fitted with Time points and Protocol as random interaction. However, no difference was found in terms of the number of probes with fold changes either >2 and/or <1.5. There were also lower numbers of probes achieving statistical significance with a random interaction.

[1] estimated using a uniform-beta mixture distribution of significance probabilities (Allison et al., 2002)
[2] estimated using Student's t distribution

A further analysis of the data was performed assuming a fixed interaction between Time point and Protocol. For the most part, significance probabilities of effects determined via the random interaction model are conservative at lower values of probability compared with those of the average effects in the presence of a fixed interaction. We chose to move forward with the output from the analysis with the fixed interaction. The reasons for considering the time point and protocol interaction in microarray statistical analysis was due to the complexity of the experimental design for this study. The complexity of the experimental design includes multiple brushing given to the same sheep at different time points and each of the sheep was subjected to the different combination of the time points. Introduction of these two components (timepoints and protocols) during the experiment, has indirectly introduced another sources of variation which might be contributed to the changes in expression profiles observed. Therefore, these two sources of variation cannot be ignored in which preliminary analysis revealed that protocols and time points contributed to the expression of the probes. This was also supported by the morphometric analysis and also preliminary experimental design for Protocols A and B in which even though the proliferative response in both protocols were pretty much similar, at specific time points (days 1 and 3) some variations (differences) were observed. Significance probabilities (for all probes at different time points following injury vs naive airway (time point 0)) were calculated approximately using an F distribution with degrees of freedom 2 and 12. The proportion of probes with such interactions was estimated by a Bonferroni method.

In this analysis, the numbers of probes with significant changes in expression at the given *p*-values were determined (Table 5.8). Across all time points there was a total of 13,495 probes that were significantly changed at the $p<0.05$ level. The highest number of changes was at 6 hours and the lowest at day 7. Amongst the 13,495 probes, there are 6,420 probes been annotated (details of the annotation process will be discussed in the Chapter 6). These 6,420 annotated probes were used for downstream application for the gene functional and pathway analyses. The list of 6,420 probes and the details of statistical outputs is shown in the (Chapter 6_ Appendix 6.1 on CD).

Time point (vs. day 0)	% of probes differentially expressed[1]	Number of probes with fold change		Number of probes statistically significant[2]		
		≥2	<1/2	*p*<0.05	*p*<0.01	*p*<0.001
6 hours	49	**252**	1042	**4749**	2424	639
1 day	34	**425**	332	**2785**	1167	366
3 days	43	**582**	363	**3556**	1500	446
7 days	37	**342**	155	**2405**	885	254

Table 5.8: Statistical analysis of microarray data using a linear mixed model with a fixed interaction between the time points and protocols.
Table illustrates the numbers of probes that are differentially expressed at each time point either by fold change greater than 2 up or down or by *p*-value.

5.4.2.4 Validation of microarray study

5.4.2.4.1 Ki67 vs PCNA

Further molecular analysis was performed to validate the array data for the proliferating cell nuclear antigen (PCNA) gene expression with Ki67 staining data – both of which are involved in proliferative activity. In the array experiment, the PCNA gene was highly significantly change at day 3 post injury (p=0.002) with fold change of 3.23 whilst no change at day 1 ($p<0.001$; FC=1.0) when compared to the naïve airway (time point 0). This compares favourably with the morphometric analysis of Ki67-stained cells where proliferative activity was also significantly increased ($p<0.05$) from day 1 to day 7 post injury when compared to naive.

In order to compare these two datasets, the Ki67 counts for damaged and transition areas from treated airways were combined since this is the equivalent region harvested for the gene expression study. The dataset for naïve control airways (time point 0) were included. Statistical analysis was performed to assess the correlation between PCNA array data and positive Ki67 cell counts (equivalent bronchial brushing sites from the same sheep at same time point). A correlation coefficient test was performed using Pearson with $p<0.05$ considered significant. The data shows a correlation between PCNA expression and Ki67-positive cells ($r = 0.45$, $p = 0.008$) – (Figure 5.10).

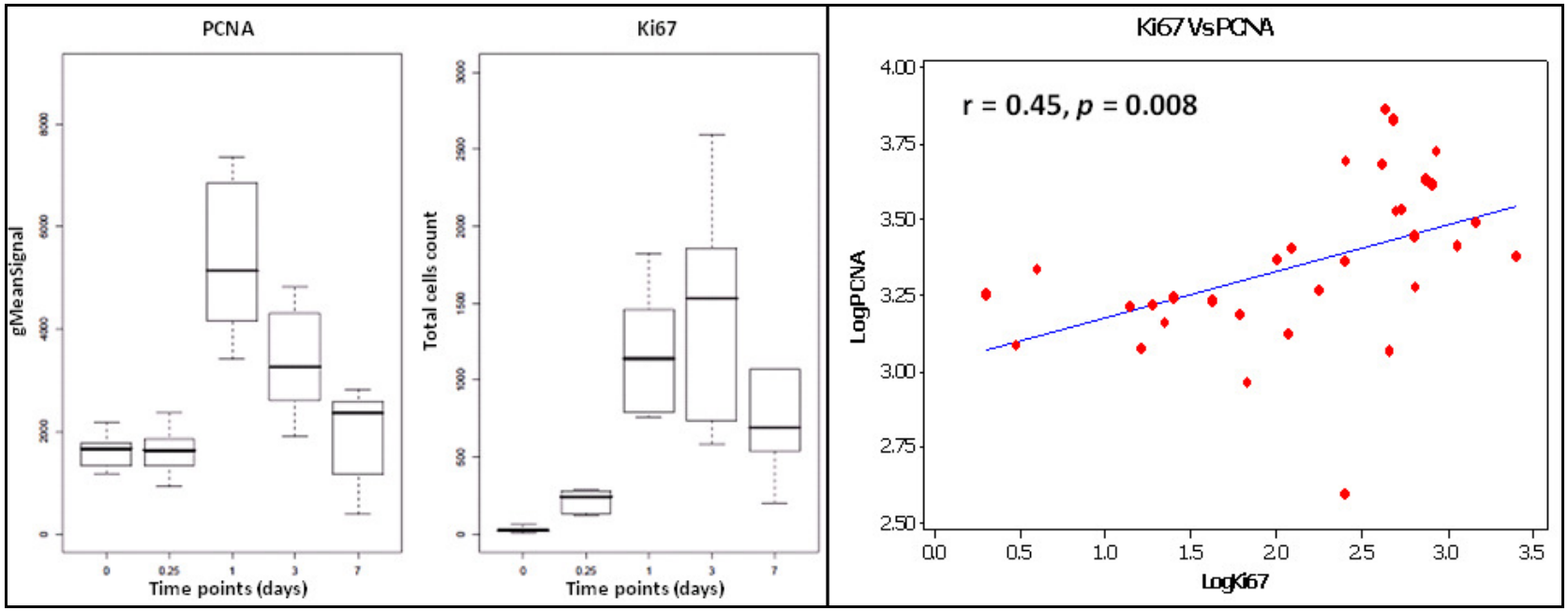


Figure 5.10: A comparison of the two indicators of cell proliferation
Figure illustrates a) the profiles of PCNA expression (p<0.05 compared to naive at days 1 and 3) and Ki67 positive cell counts (p<0.05 at all time points except 6 hours) over time and b) Correlation between paired Ki67 counts and PCNA expression levels (r = 0.45, p = 0.008). Box plots on the left panels for PCNA were plotted based on the mean values of green signal density (gMeanSignal) whilst the Ki67 was plotted based on the absolute numbers Ki67-positive cells counted during morphometric analysis (Chapter 4; Section 4.4.5 –damaged and transitional areas). To generate the correlation between Ki67 vs PCNA, the count data for Ki67 staining for damaged and transitional areas (Chapter 4; Section 4.4.5) corresponding to the areas used for microarray experiment were used. Upper and lower box plot margins represent the interquartile range; middle bar indicates the median. Scatter plot on the right panel was plotted based on the data from the left panels which were logged-transformed.

5.4.2.4.2 AB-PAS staining vs mucus-related genes

This analysis was performed to validate the array data for genes related to mucus-production with the AB-PAS positive cell staining data. Although there are many genes associate with the mucus production in the airway, for the purpose of this study, we only selected genes from the array data for interleukin 4 (IL4) and interleukin 13 (IL13) based on the fact that these genes play a role in mucus production in response to injury and repair (Dabbagh K et al., 1999, Shim JJ et al., 2001). Expression of these genes was unchanged throughout the times following brushing (Table 5.9). The expression patterns for both genes can be seen in the (Figure 5.11). From the array experiment, it shows that the IL4 gene was significantly unchanged at day 1 and day 7 post injury (FC at 0.94 and 0.93, respectively). The IL13 gene was also significantly unchanged at day 3 post injury (FC at 1.16). Expression of other genes that relate to the mucus production such as EGFR and TNF-alpha were not significantly altered.

In order to assess whether the distribution of AB-PAS positive stained cells correlated with the expression of these genes, the AB-PAS positive cell count data from treated airways were adjusted to include only the data for damaged and transition areas for each segment since this is the equivalent region harvested for the gene expression study. However, only the undamaged data for naïve control segments were included. The absolute numbers of AB-PAS positive cells count and gMeanSignal were both normally distributed. As described for the Ki67 staining dataset, the data were transcribed according to the sheep ID, time points and the number of AB-PAS stained counted. Statistical analysis was performed by using a Pearson correlation coefficient test with $p<0.05$ was considered significant. Neither the expression of these genes had significant correlation with the number of goblet cells on the correspond areas of measured (damaged and transitional) – IL4 ($r=0.13$, $p=0.4772$) and IL13 ($r=0.30$, $p=0.0846$) – (Figure 5.12).

Time points (days)	0.25		1		3		7	
Genes	**IL4**	**IL13**	**IL4**	**IL13**	**IL4**	**IL13**	**IL4**	**IL13**
FC	0.93	1.1	0.92	1.1	0.97	1.1	0.93	1.0
***p*-value**	0.08	0.05	0.01	0.13	0.37	0.03	0.02	0.20

Table 5.9: The significant probabilities and fold changes of the IL4 and IL13 genes over the time of injury
The IL4 gene was significant changed ($p<0.05$) at day 1 and day 7 post injury whilst IL13 gene was only significantly changed ($p<0.05$) at day 3 post injury.

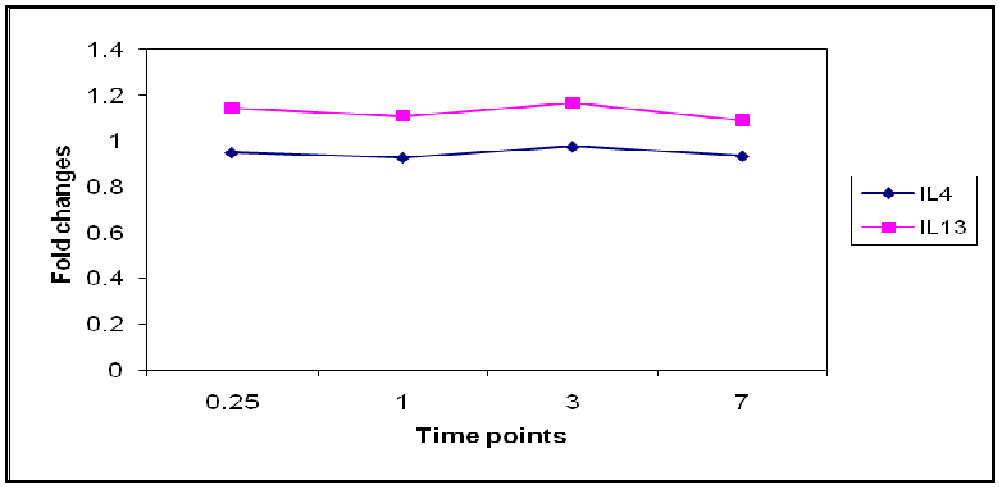


Figure 5.11: The expression patterns of the IL4 and 13 genes based on the fold changes across the time of injuries. A similar data was taken from (Table 5.9).
The significance changes observed from statistical analysis of microarray data are considered up- and/or down-regulated based on the values of fold changes between <0.5 (down-regulated), >2.0 (up-regulated) and 0.5>FC<2.0 (unchanged).

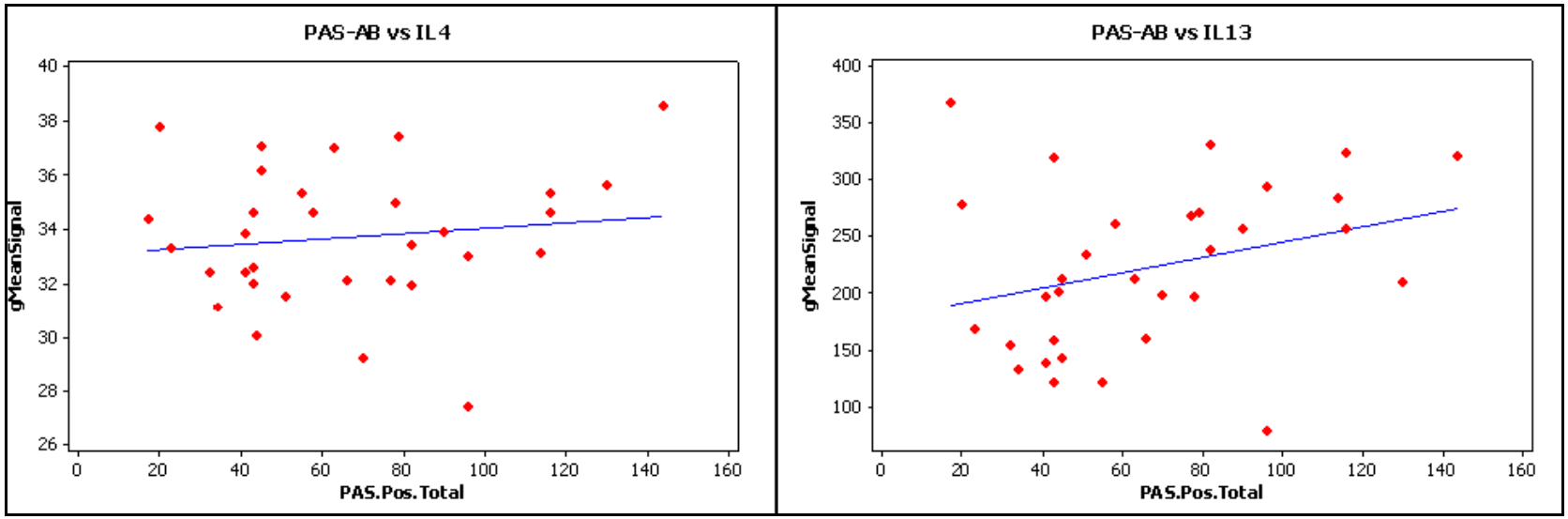


Figure 5.12: An association between interleukin genes IL4 and IL13 plotted against number of AB-PAS stained cells.
No correlation was observed for expression of either gene (p>0.05) with the distribution of AB-PAS stained cells.

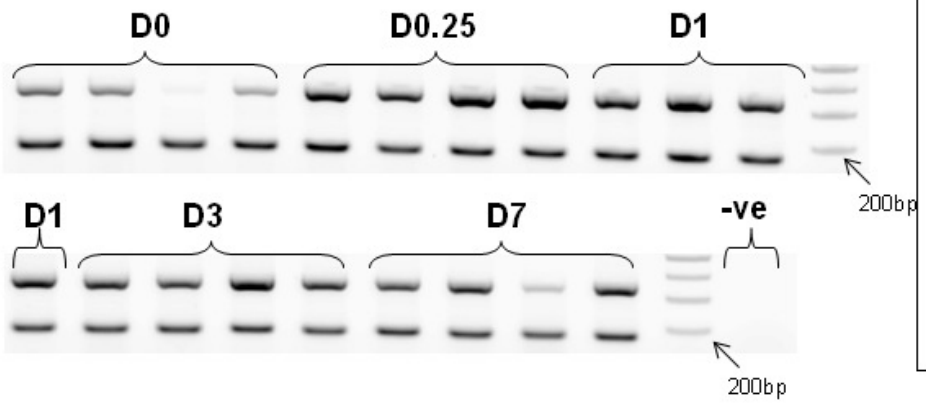
5.4.2.4.3 Semi-quantitative RT-PCR analysis

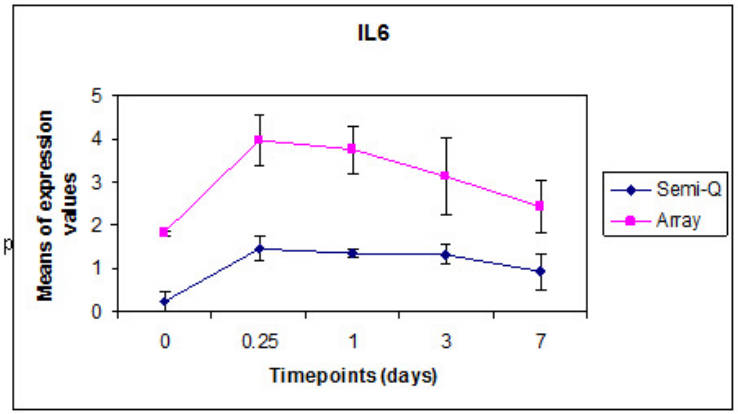
Semi-quantitative RT-PCR assay was performed to validate 9 genes including IL6, keratin 19 (KRT19), keratin 6A (KRT6A), cell division cycle 2 (CDC2), cyclin B1 (CCNB1), Cyclin B2 (CCNB2), keratin 5 (KRT5), tubulin beta (TUBB) and catenin beta-like 1 (CTNNBL1) – (Table 5.10). All the PCR conditions (cycles and temperature) for both genes of interest and internal control (ATPase) were successfully optimised. The RT-PCR assays were developed and the genes of interest successfully amplified. The amplified signals were quantified from agarose gels relative to the expression of the ATPase gene. The expression values detected by RT-PCR assay for these genes were compared with the gMeanSignal expression values detected by the array. Majority of the samples (for array dataset) were Log-transformed to the LoggMeanSignal since the data were not normally distributed. The expression profiles were plotted as a mean value (±SD) for each array (n=4 for each time point – data equivalent to the dataset for RT-PCR) and semi-quantitative RT-PCR (n=4 for each time point) data – (Figure 5.13). Pearson correlation coefficient tests showed that the expression values detected by the semi-quantitative RT-PCR assay had significant correlation with the expression values detected by the array ($p<0.05$) except for the beta-catenin gene – (Figure 5.14).

Genes	Fold change				*p*-value (interaction)			
	Day 0.25	**Day 1**	**Day 3**	**Day 7**	**Day 0.25**	**Day 1**	**Day 3**	**Day 7**
KRT6A	24.30	55.51	75.87	43.47	*p*<0.001	*p*<0.001	*p*<0.001	*p*<0.001
IL6	54.40	54.72	23.65	4.39	*p*<0.001	*p*<0.001	*p*<0.001	0.028909
KRT19	7.33	16.99	20.43	13.50	*p*<0.001	*p*<0.001	*p*<0.001	*p*<0.001
CDC2	0.73	5.29	7.39	2.47	0.218	*p*<0.001	*p*<0.001	0.002193
CCNB1	0.82	7.78	7.78	2.87	0.254	*p*<0.001	*p*<0.001	*p*<0.001
CCNB2	1.05	3.77	4.56	2.21	0.815	*p*<0.001	*p*<0.001	0.002
KRT5	1.50	2.40	2.50	1.79	0.059	*p*<0.001	*p*<0.001	0.011
TUBB	1.15	2.14	1.48	1.56	0.148	*p*<0.001	0.001	*p*<0.001
CTNNBL1	0.90	1.14	0.92	0.73	0.557	0.426	0.639	0.091

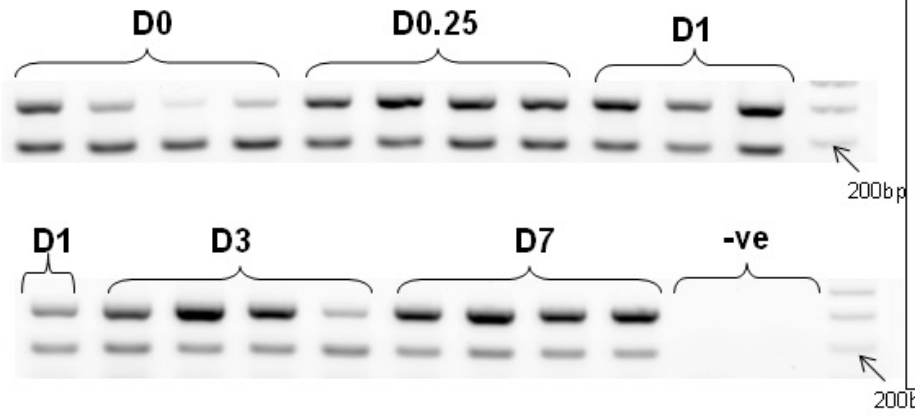
Table 5.10: A list of genes subjected for a validation study using a semi-quantitative RT-PCR assay
A significant probability and fold changes observed for each time point (n=32 samples in total) calculated using a linear mixed model relative to naive control by taking into consideration of fixed interaction between the time points and the protocols.

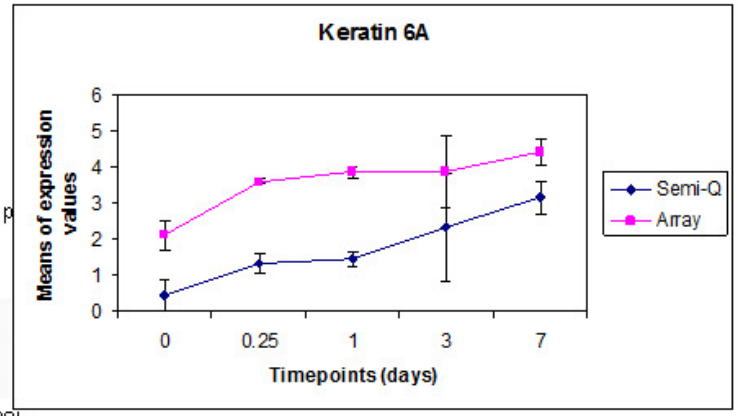
IL6



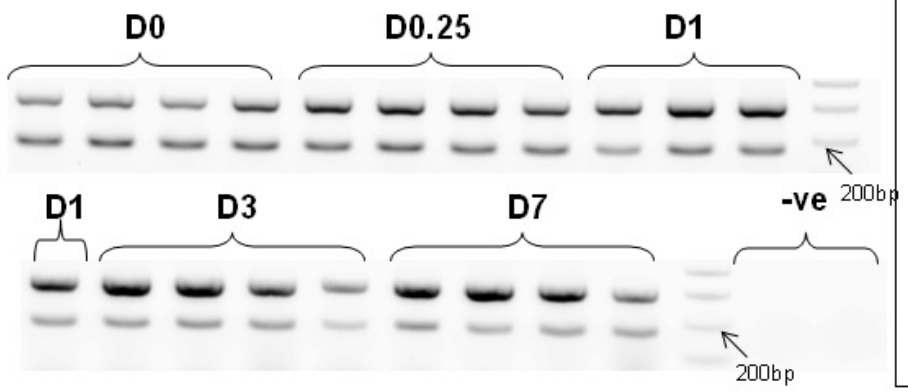


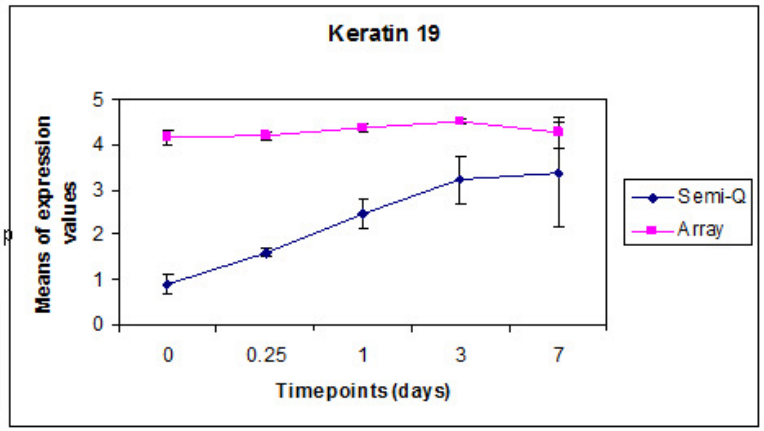
Keratin 6A



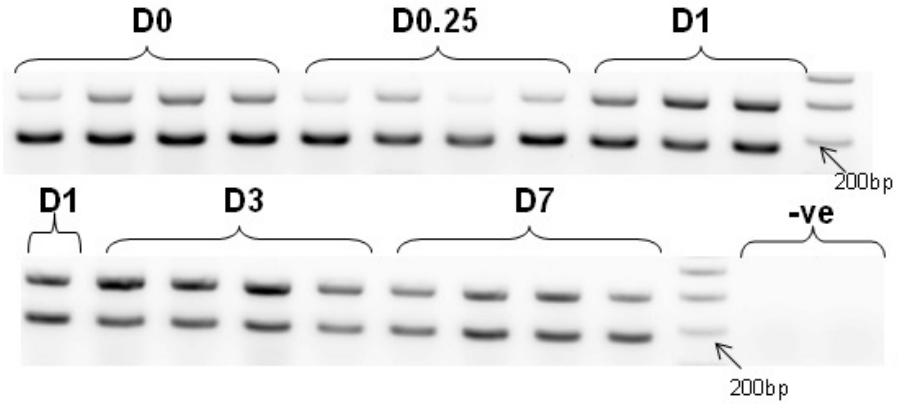


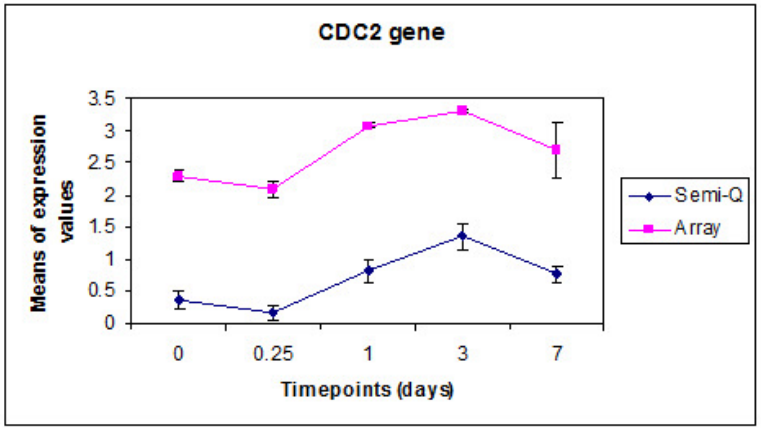
Keratin 19



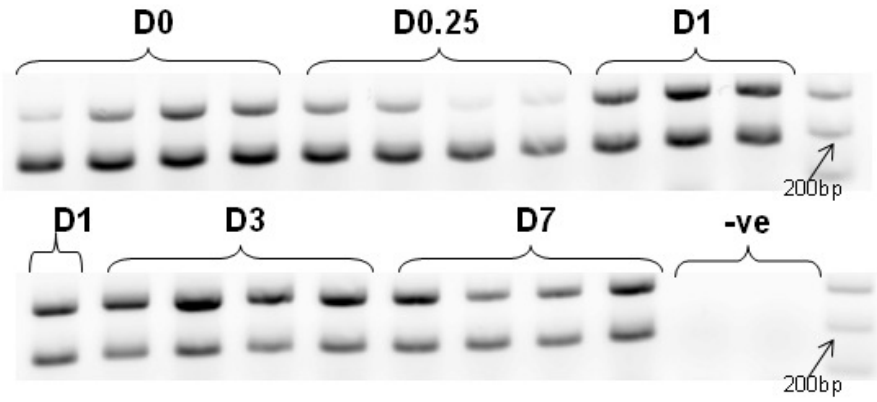


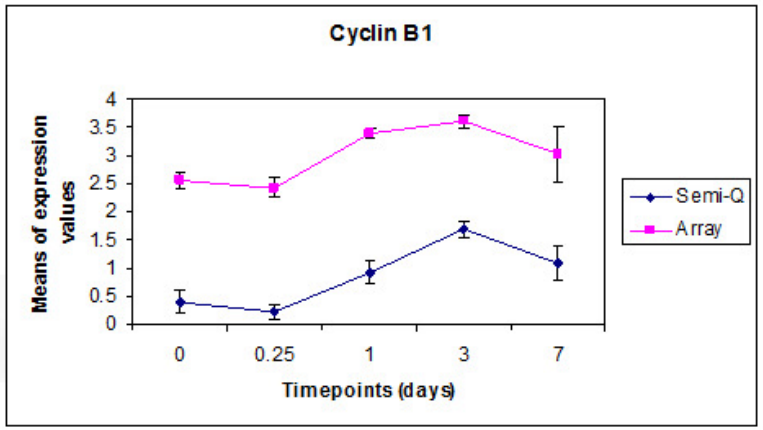
CDC2



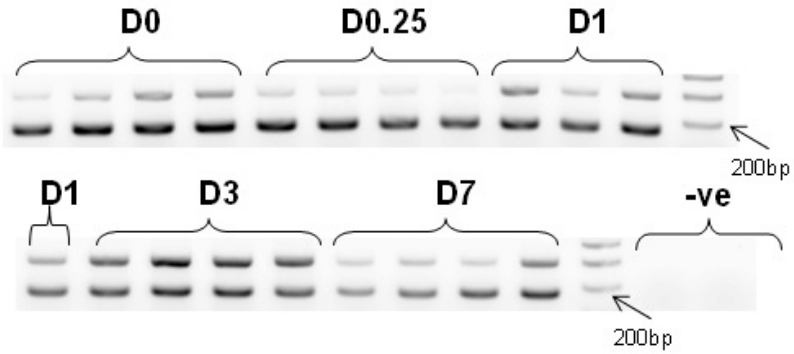


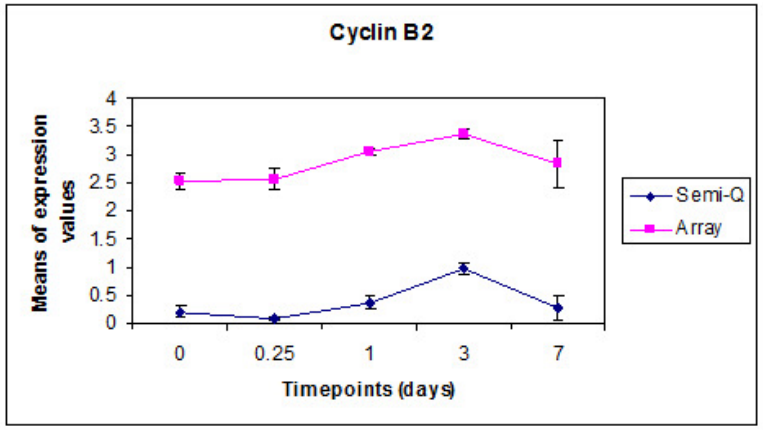
Cyclin B1



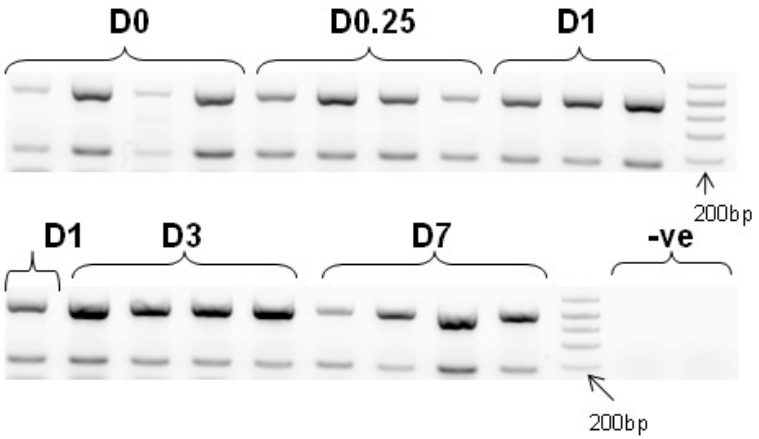


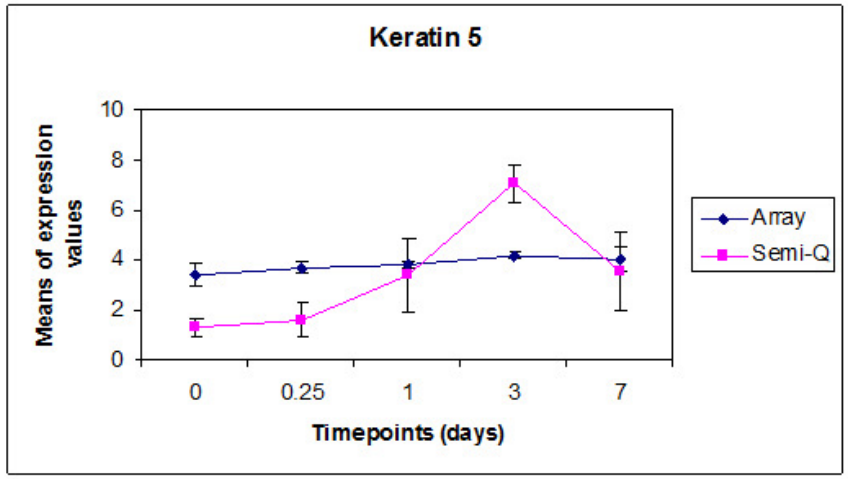
Cyclin B2



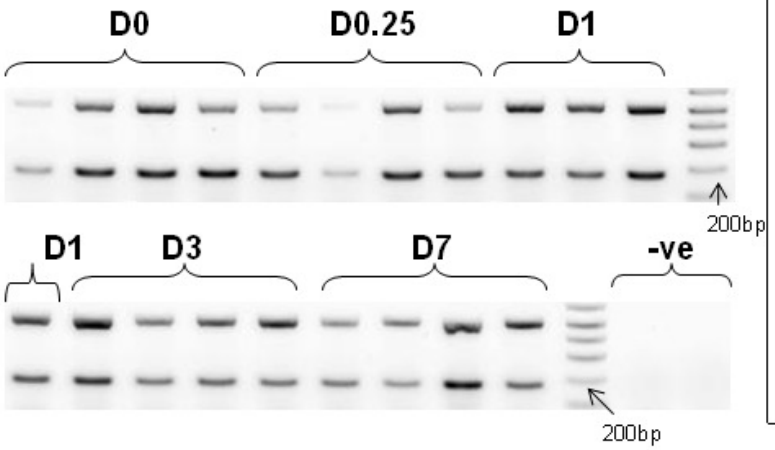


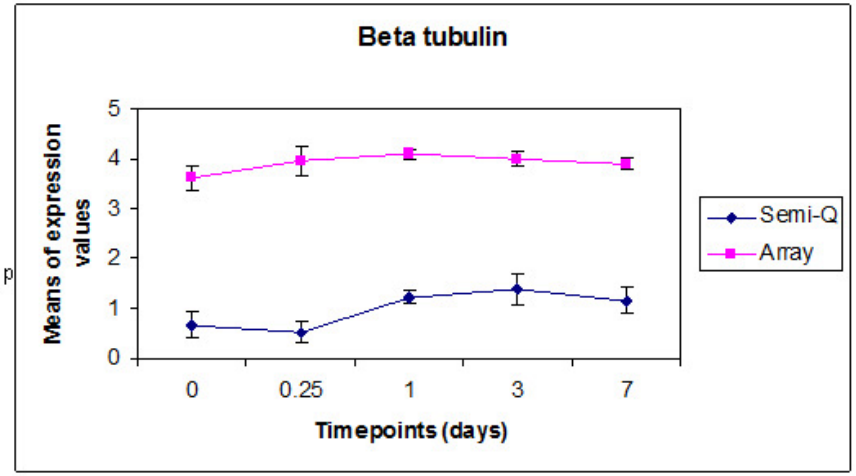
Keratin 5





Beta tubulin





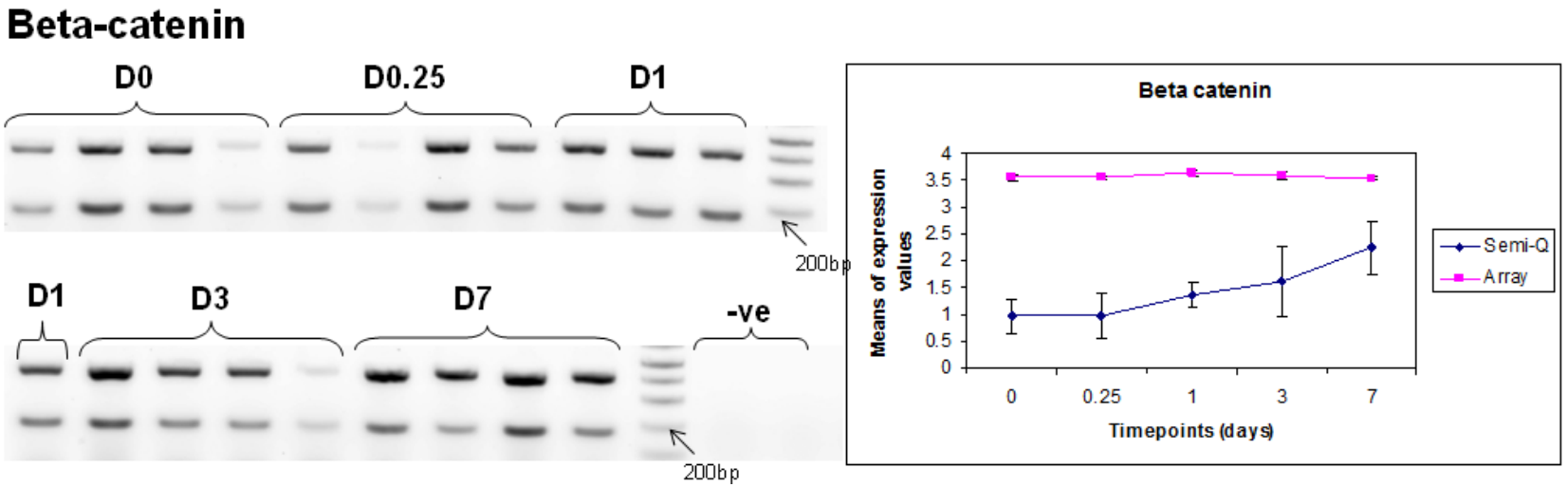


Figure 5.13: Semi-quantitative RT-PCR expression profiles compared to Microarray data for selected genes.
RT-PCR analysis for genes of interest (left panels). A 100bp marker was used to estimate the DNA sizes. The PCR products at the bottoms are the product for ATPase whilst at the upper lines are the products for the genes of interest. The mean values for both data – RT-PCR (semi-Q) and gMeanSignal (Array) were plotted across the time of injuries. Most of the gMeanSignal for array dataset were Log transformed since not normally distributed. X-axis shows the time points of where the samples were collected following brushing and time point 0 refers to the samples taken from naive control segments, and the Y-axis shows the mean values of four samples for each time point, for both RT-PCR assay (relative to the ATPase) and array (intensity of green fluorescence). (* D = day, -ve = PCR negative control with ddH_2O as a template)

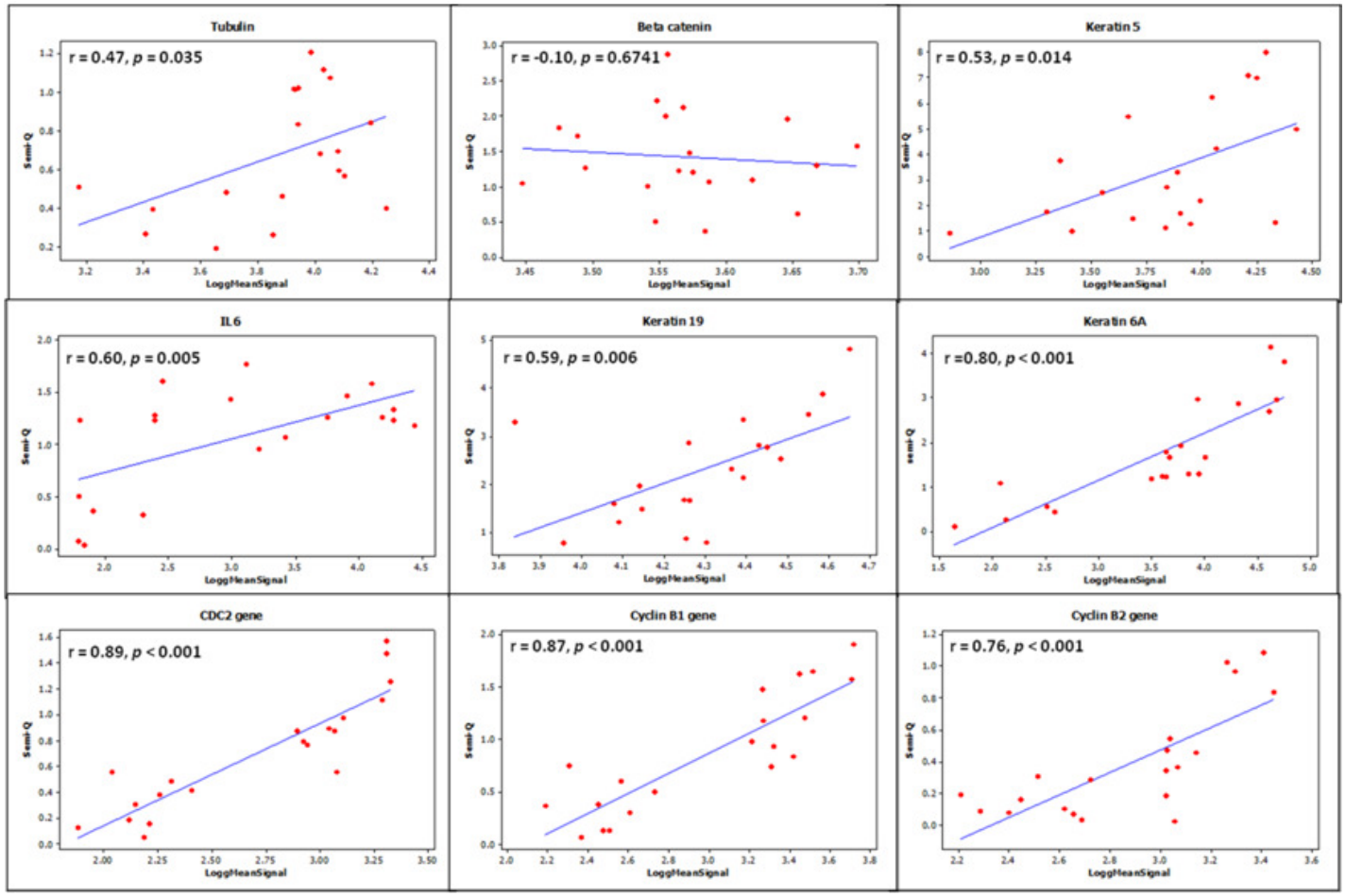


Figure 5.14: A correlation coefficient test between the data from microarray (gMeanSignal) and a semi-quantitative RT-PCR assay (relative to the ATPase) for the genes subjected for validation study.
The X-axis shows the log transformed array data and the Y-axis shows the semi-quantitative RT-PCR for each sample (n=4 for each time point – relative to the ATPase). All genes show a positive correlation (p<0.05) between microarray data and a semi-quantitative RT-PCR assay for the selected samples except for the CTNNBL1 gene. (* Semi-Q = semi-quantitative RT-PCR)

In order to understand the patterns of genes expression regulation in all annotated probes which significantly expressed at $p<0.05$, the average effects were plotted according to the time points (Figure 5.15). This figure shows that the numbers of probes been up-regulated were increased over the time following injury in which most of the significantly altered probes at 6 hours were down-regulated.

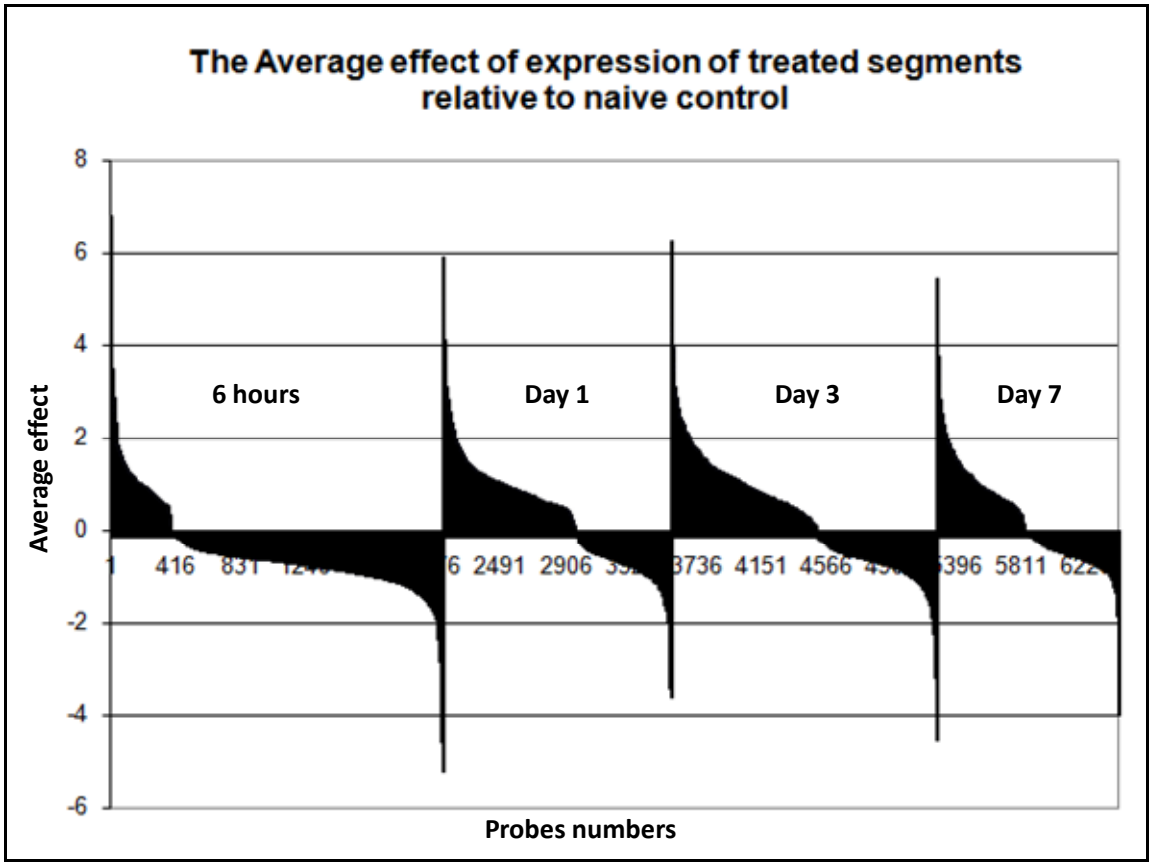


Figure 5.15: The profiles of the significantly regulated probes following bronchial brushing
This figure was plotted based on the average effect values obtained from the statistical analysis for all annotated probes that significantly expressed following injury. The x-axis shows the list of the probes and y-axis shows the average effect of the probes which are listed from highest to lowest values in each time points. The average effect values plotted in this graph is taken from the data presented in the Appendix 6.1 (on CD) of "effect" in the coumn H. Since the values for the average effect reflects the fold changes of the probes, by looking at this figure one could expect to have the high fold changes with the increase of average effect values. The chages of the probes considered up-regulated if the average effect >1, down-regulated < -1 whilst no changed if -1 > average effect <1. The 6 hours time point shows a massive numbers of genes been regulated following a bronchial brushing, however majority of the regulated genes at this time point are down-regulated and unchanged. The numbers of up-regulated genes are gradually changed over time and by day 7 the numbers of both up- and down-regulated genes are pretty much similar.

Due to the massive numbers of genes that being significantly altered following bronchial brushing, it is impossible to highlight all these genes in this chapter. One way of presenting this dataset is to extract the genes with most robust changes in both up- and down-regulated. A number of genes that exhibited greater than 10-fold increases in expression were identified (Table 5.11). Most of the genes that were highly up-regulated are involved in cytokine activity, immune and inflammatory responses and response to wound such as metallothionein 1E (MT1E), chemokine (CXC motif) ligand 5 (CXCL5), lipopolysaccharide binding protein (LBP), pentraxin 3 (PTX3), interleukin 6 (IL6), selectin P (SELP) and interleukin 8 (IL8), cellular proliferation and differentiation [keratin 6A (KRT6A), keratin 19 (KRT19), ADAM metallopeptidase domain 12 (ADAM-12)], and cellular migration and adhesion [thrombospondin-2 (THBS2), olfactomedin 4 (OLFM4), SELP, ADAM12 and IL8). Amongst those down-regulated changes were pulmonary surfactant-associated protein A2B (SFTPA2B), surfactant protein C (SFTPC), CD5 molecule-like (CD5L), complement component (3d/Epstein Barr virus) receptor 2 (CR2), macrophage receptor with collagenous structure (MARCO) and cytochrome P4501A1 (CYP1A1) – (Table 5.12).

Probenames	TargetID	GeneSymbol	GeneName	Fold change			
				0.25	**1**	**3**	**7**
A_70_P021596	DY511316	ADAM12	ADAM metallopeptidase domain 12	2.26	3.77	**11.08**	**16.36**
A_70_P054711	EE770225	AGP	Alpha-1 acid glycoprotein	1.16	1.93	8.20	**13.60**
A_70_P030361	NM_001009366	BLG	Beta-lactoglobulin	6.76	**15.17**	**15.64**	3.46
A_70_P002816	AY392761	CHI3L1	Signal processing protein	5.00	**11.04**	5.40	3.53
A_70_P001891	EE774505	COL1A1	Collagen, type I, alpha 1	1.17	1.96	**10.81**	8.25
A_70_P059886	DY496997	COL1A1	Collagen, type I, alpha 1	1.24	1.44	9.38	**13.35**
A_70_P030126	NM_001009733	CP	Ceruloplasmin	6.17	**24.16**	**13.97**	3.76
A_70_P030127	NM_001009733	CP	Ceruloplasmin	8.34	**32.22**	**19.14**	4.86
A_70_P058706	FE033277	CXCL5	Chemokine (C-X-C motif) ligand 5 (CXCL5)	**11.25**	**14.56**	**18.55**	6.17
A_70_P038041	EE791827	HP	Haptoglobin (HP)	**19.55**	**32.51**	**12.05**	5.54
A_70_P039706	NM_001009392	IL6	Interleukin 6	**54.40**	**54.72**	**23.65**	4.39
A_70_P039661	NM_001009401	IL8	Interleukin 8	**10.74**	6.06	**27.23**	8.40
A_70_P039662	NM_001009401	IL8	Interleukin 8	9.16	5.30	**25.35**	7.50
A_70_P002596	EE860756	KRT19	Keratin 19 (KRT19)	7.33	**16.99**	**20.43**	**13.50**
A_70_P068966	EE837811	KRT6A	Keratin 6A (KRT6A)	**24.30**	**55.51**	**75.87**	**43.47**
A_70_P045231	EE787299	LBP	Lipopolysaccharide binding protein (LBP)	**44.82**	**49.20**	**13.41**	4.52
A_70_P026536	AF267156	MMP-1	Interstitial collagenase	3.50	3.67	**20.74**	3.20
A_70_P026761	DY513601	MMP-13	Matrix metalloproteinase 13	1.40	1.35	**15.37**	**10.84**
A_70_P037411	EE798177	MT1E	Metallothionein 1E (MT1E)	**110.22**	**59.84**	**15.15**	2.27
A_70_P009666	EE765361	MT-2	Metallothionein 2A	**13.99**	**14.71**	**11.74**	2.56
A_70_P064826	EE851232	OLFM4	Olfactomedin 4 (OLFM4)	**14.21**	**22.43**	4.82	1.39
A_70_P010836	EE829599	PTX3	Pentraxin-related gene, rapidly induced by IL-1 beta	**10.79**	6.92	2.75	1.85

A_70_P004271	EE804747	RRM2	Ribonucleotide reductase M2 polypeptide transcript variant 3 (RRM2)	0.60	**13.80**	9.23	4.82
A_70_P019681	EE866720	SAA3	Serum amyloid A 3 (SAA3)	**14.96**	**20.68**	**14.74**	9.67
A_70_P039517	NM_001009295	SELP	Selectin P (granule membrane protein 140kDa, antigen CD62)	**10.97**	**12.06**	4.47	2.09
A_70_P060116	EE772499	THBS2	Thrombospondin 2	1.81	2.03	5.09	**10.53**
A_70_P038291	EE830791			2.02	2.82	7.77	**10.09**

Table 5.11: The list of significantly up-regulated genes with the FC more than 10 at any given time of injury (in bold)
Most of these genes are involved in inflammatory response, innate immunity, cellular migration, proliferation and differentiation.

Probenames	TargetID	GeneSymbols	GeneName	Fold change			
				6 hours	Day 1	Day 3	Day 7
A_70_P051201	NM_001009721	GHRL	ghrelin/obestatin prepropeptide	0.451277	0.338967	0.231581	**0.07697**
A_70_P051202	NM_001009721	GHRL	ghrelin/obestatin prepropeptide	0.390526	0.317788	0.193863	**0.062363**
A_70_P045486	EE862425	CYP2B	cytochrome P450 subfamily 2B	0.191148	0.217368	**0.074232**	0.42846
A_70_P019746	EE788235	HSD17B11	hydroxysteroid (17-beta) dehydrogenase 11	0.121486	0.216232	**0.092625**	0.612878
A_70_P019656	EE789897	ADH1C	alcohol dehydrogenase 1C (class I), gamma polypeptide	0.103437	0.10427	**0.072873**	0.346057
A_70_P030377	NM_001009189	SHEUGT1A07	UDP-glucuronosyltransferase	**0.097427**	0.133277	0.100579	0.452924
A_70_P045421	EE790996	BSP30C	common salivary protein BSP30 form C	**0.097156**	0.162488	0.111051	0.342399
A_70_P045421	EE790996	BSP30C	common salivary protein BSP30 form C	**0.097156**	0.162488	0.111051	0.342399
A_70_P003891	FE033669	CLDN18	claudin 18	**0.084601**	0.126006	**0.08729**	0.410856
A_70_P041841	S79795	CYP1A1	cytochrome P4501A1	**0.074732**	0.110169	**0.059738**	0.300586
A_70_P020236	EE786153	BSP30C	common salivary protein BSP30 form C	**0.071274**	0.132524	**0.086807**	0.34502
A_70_P039392	NM_001009724	CR2	complement component (3d/Epstein Barr virus) receptor 2	**0.064773**	0.13942	0.183443	0.579722
A_70_P045261	EE825133	MARCO	macrophage receptor with collagenous structure	**0.051556**	**0.092943**	**0.043158**	0.427027
A_70_P039391	NM_001009724	CR2	complement component (3d/Epstein Barr virus) receptor 2	**0.04813**	0.118102	0.17051	0.429781
A_70_P014851	CN821988	CD5L	CD5 molecule-like	**0.043726**	**0.089881**	**0.057625**	0.431331
A_70_P039301	NM_001009729	SFTPC	surfactant protein C	**0.041848**	**0.0998**	0.128709	0.260685
A_70_P051276	NM_001009728	SFTPA2B	pulmonary surfactant-associated protein A	**0.030921**	0.125114	**0.079272**	0.215829
A_70_P039302	NM_001009729	SFTPC	surfactant protein C	**0.030805**	**0.081981**	**0.09456**	0.225396
A_70_P051277	NM_001009728	SFTPA2B	pulmonary surfactant-associated protein A	**0.026759**	0.105296	**0.068933**	0.20601

Table 5.12: List of significantly down-regulated genes with the FC more than 10 at any given time of injury (in bold)
Majority of these genes are involved in regulation of liquid surface tension, respiratory gaseous exchange, homeostasis process and defence response. There are more than one probes encode for the same genes.

5.5 DISCUSSION

The main objectives of the study presented in this chapter were to optimise the integrity and quality of the RNA samples isolated from the cartilaginous airway tissues for the microarray study, and to investigate the expression profiles involved in airway regeneration and repair. RNA yield as well as quality and integrity are crucial elements for success and reliability of microarray-based expression profiling for thousands of genes in a single assay. However, a lower RNA yield is expected in the lung tissue compared to other solid organs (Fleige S et al., 2006) due to the presence of cartilage tissue in the airway. Cartilage in airway samples presents potential problems in RNA isolation as cartilage has low cell content and contains an extracellular matrix rich in proteoglycans, which tend to co-purify with the RNA (Fleige S et al., 2006).

In the present study, quality of the isolated RNA was initially assessed by conventional UV spectrophotometric absorbance. The A_{260}/A_{280} and A_{260}/A_{230} ratios reflected RNA purity, but were not informative regarding the integrity of the RNA (Imbeaud S et al., 2005). Low A_{260}/A_{280} ratios are indicative of protein contamination, whereas low A_{260}/A_{230} ratios serve to indicate the absence of residual contamination by organic compounds such as phenol, sugar or alcohol, which could be highly detrimental to downstream applications (Imbeaud S et al., 2005). Sambrook and colleagues (1989) suggested that a A_{260}/A_{280} ratio of greater than 1.8 indicates a good quality of the RNA (Sambrook J et al., 1989). However, it was reported that samples with A_{260}/A_{230} ratios <1.8 did not exhibit any inhibition during downstream applications such as cDNA synthesis and labelling or *in vitro* transcription (Imbeaud S et al., 2005).

The current accepted standard for expressing information pertaining to RNA quality and integrity is the RNA integrity number (RIN). The RIN score is generated as part of the Agilent 2100 bioanalyzer platform and represents the results of an algorithm applied to the electropherograms (Schroeder A et al., 2006). Thus the quality and integrity of the samples assessed in this study was not determined simply by the ratio of the ribosomal bands but determined by the entire profile of the electrophoretic traces of the RNA. The

RIN therefore takes into consideration the ratio of the 28S:18S ribosomal RNA peaks,.the appearance of additional peaks and elevation of the baseline in the region between ribosomal peaks (Copois V et al., 2007). Indeed, in relation to the 28S:18S rRNA ratios, a study by Imbeaud S et al., (2005) cast doubt on their value as a gold standard for assessing RNA integrity. Whilst these authors found that there was no correlation between 28S:18S ratios and RNA integrity, RNAs with 28S:18S >2.0 were usually of high quality and 28S:18S ratios >1.0 were reported to be of good quality (Imbeaud S et al., 2005).

In the context of our preliminary study of optimising the RNA preparation for the microarray experiment, the integrity of the RNA was assessed using the RIN score. Initial RNA samples from 30mg cartilaginous airway tissue failed to generate good RIN values although the average of the A_{260}/A_{280} ratio for all samples was greater than 1.8 (Table 5.3 and Figure 5.4). This study demonstrated that there was no association between A_{260}/A_{280} ratio and RIN score (Figure 5.5; right panel). Increasing the amount of airway tissue used in RNA preparation to 100mg generated acceptable RIN values for the majority of samples although some of the RIN scores were still below the Agilent quality control standard (RIN $\geq$7). Whilst the profiles of the two ribosomal bands can be seen very clearly in those samples with RIN below 7 (particularly those with RIN between 6 and 7), the profiles also reflect the appearance of additional peaks before the 18S peak – which may be due to the presence of short fragments of degraded RNA – (Figure 5.3). The height of the 28S peak contributes additional information about the state of the degradation process since the 28S band disappears faster than the 18S band thus allowing early detection of the degradation process (Schroeder A et al., 2006). As we were unable to consistently achieve a RIN $\geq$7 with the cartilaginous airway tissue, an internal compromise on this target was forced, and RINs in excess of 6.0 were considered acceptable as long as both ribosomal peaks were clearly observed and their presence not obscured by additional peaks before and between these ribosomal bands. The low RIN score may be a feature of dealing with RNA from lung tissue samples. Fleige S et al., (2006) assessed RIN scores from various bovine tissue samples and

reported that the average RIN score for lung tissue was 6.55±0.67 (Fleige S et al., 2006, Fleige S and Pfaffl MW, 2006) whilst RIN scores as high as 9.62±0.32 were obtained from Corpus luteum (Fleige S and Pfaffl MW, 2006).

Low RIN scores normally associate with degraded RNA samples and such degradation can often cause a problem when handling RNA samples extracted from formalin-fixed, paraffin-embedded (FFPE) tissue (Farragher SM et al., 2008). However, a study by (Roberts L et al., 2009) reported that microarray data generated from FFPE samples from lung and colon tissues were comparable with data from frozen samples. Whilst the RIN scores in the former instance were relatively low (RIN> 4) distinct peaks of 18S and 28S rRNA were observed. This study demonstrated that such samples (clear 18S and 28S ribosomal RNA peaks and RIN>4) can reliably generate microarray data with greater than 35% of the probe sets detected as present and indeed would usually generate sets with greater than 50% detected. Conversely significantly degraded and modified samples, those without distinct 18S and 28S peaks, had less than 35% probe sets detected as present on the microarray (Roberts L et al., 2009).

The present study revealed that a larger amount of starting material of the airway samples gave an increased concentration of RNA and generally resulted in improved RNA integrity. However, a more robust analysis involving the samples for the microarray study showed that there was no correlation (p>0.05) between RNA concentration or A260:A280 ratio and RIN score. After several optimization strategies failed to consistently achieve improved quality and integrity of RNA, concerns were raised that this might be an issue specific to extraction of RNA from cartilaginous airway tissues and hence could negatively impact on our strategy of following molecular changes in the airway wall using such samples. To assess this issue we compared the RNA quality obtained from different lung tissue types (cartilaginous airway vs parenchyma vs bronchial brushing). We observed that the quality of RNA extracted from cartilaginous airways was consistently poorer than that obtained from other lung-derived samples. We also observed that the RIN values of RNA isolated from

parenchymal samples were of optimal quality (>7). Notably, despite the relatively poor yields of RNA obtained from the small bronchial brushing samples used in this study, the quality of RNA isolated from such samples was considered excellent. This concurs with reports describing the transcriptomic analyses of bronchial brushing samples obtained from lung transplant patients where such samples, which contain a relatively pure cell population of >90% bronchial epithelial cells (Skawran B et al., 2009), generated high quality RNA for microarray analysis (Raman T et al., 2009, Skawran B et al., 2009). However, in the context of studying changes in the cellular and molecular aspects of airway repair, such samples will potentially not reflect the contribution of non-epithelial compartments (such as glands and smooth muscle) which may be important in the process of airway repair.

The findings from this preliminary study were used as a baseline for the microarray study in relation to obtaining optimal RNA integrity from the cartilaginous airways samples. Although the majority (> 90%) of the samples have good RNA quality and integrity scores, some of the RNA samples with a high RIN (≥7.5) failed to generate the optimal sizes of the amplified RNA (aRNA) during the amplification process. The presence of short fragment sizes might be due to the presence of an inhibitor in the RNA samples that blocked the whole amplification process. The bottle-neck appeared to be at the reverse transcription stage since a combination of increased amount of enzyme and incubation time was sufficient to overcome the problem. There are several types of inhibitors that could have been present in our samples. One of the 'classics' is lithium chloride which is commonly used as a co-precipitant for RNA precipitations and is known to inhibit RT enzymes. However, this was not the case in our samples since we used a linear acrylamide to deal with the low amounts of RNA for precipitation step. Although we had to compromise on the acceptable RIN limit for a small number of RNA samples (RIN > 6 compared to the Agilent suggested RIN ≥7), the actual microarray data output suggested that the data generated from these samples were comparable with the data generated from the samples with higher RIN numbers.

In relation to the microarray data analysis, it should be recognised that there are alternative approaches to the method of analysis described here. We chose to perform tha analysis on the gMean signal dataset on the advice of our experienced statistician however there is some debate over whether the gMean or gProcessed signal datasets are the most correct for downstream application in statistical analysis of the microarray data.

In relation to the output measures from the feature extraction protocol the gMeanSignal is considered one of the basic values. This “raw” signal is calculated from the pixel intensities contained within an “inlier” region of the spot and thus does not consider intensities relating to saturated, or not "well above background" pixels. Selection of the inlier region is often likened to the use of a “cookie cutter” (American terminology) where the best (middle) part of the dough is kept for baking and the rest discarded.

To get the gProcessed value the BG subtracted gMean signal is divided by the gMultDetrendSig value. Multiplicative detrending is used to perform an error compensation to compensate the heterogeneity of signal across the array (the "dome effect", which is due to the hybridization dynamic in the hybridization chamber). This adjustment is often more useful when intensities are higher. Whenever the BG subtracted gMean signal falls below a certain threshold (that appears to be close to the BG estimate (gLocalBGInlierAve)), a surrogate value is used. Feature Extraction calculates this surrogate signal from the standard deviation (SD) of the background used for that feature. When the “No Background” option is chosen, the SD is the pixel level SD of the local background. These SDs are reported in the Feature Extraction output columns, “gBGSDUsed”

Although the overall analysis carried in this thesis was based on the gMeanSignal generated from the Agilent FE software rather the gProcessedSignal, the same approach for statistical analysis was carried out for the gProcessed signal dataset using a fixed interaction between the time points and the protocol. The statistical analysis output indicated that although gProcessedSignal gives higher numbers of probes (up to double

in some cases) with fold change <1/2 or >2 the numbers of probes statistically significant at the 5% level are lower by 3% to7% across time points. Both statistical and biological significance should be used to indicate differential expression. If a fold change of <1/2 or >2 is regarded as biologically significant then gProcessedSignal gives higher numbers of probes which are both statistically significant at 5% and have fold change <1/2 or >2 in all time effects considered. With this criterion it seems unlikely that the use of gMeanSignal rather than gProcessedSignal has over-estimated the number of differentially expressed probes. In addition to this it is however useful to consider how many probes selected via gMeanSignal would also have been selected using gProcessedSignal. If a subset of probes is selected for further investigation or discussion using the criterion of statistical significance and fold change mentioned above then nearly all probes (in excess of 98%) selected using gMeanSignal would also be selected using gProcessed signal. The details of the statistical analysis outputs are shown in the Supplement 11 on the CD.

The rationale for subtracting a measure of background from the signal intensity is the assumption that the fluorescence of a spot is the sum of signal intensity and some background noise which is due to a variety of technical factors such as non-specific hybridization and fluorescence emitted from other chemicals on the glass. If background is included in the estimate of the signal for a spot, then the result is a biased estimate of the true hybridization. This is particularly prominent at low signal intensities, resulting in an underestimate of the fold changes. On the other hand, subtraction of background involves an additional estimate which will increase the variability of the signal log ratios.

Background correction was not selected for in our protocols. This is consistent with a recent study of pre-processing methods appropriate for Agilent array data in which the authors stipulated a preference for no background subtraction arguing the case from a primary philosophy of avoiding unnecessary data processing but also presenting data that Agilent's pre-processing with background subtraction results in increased

variability, especially at the low intensity end (Zahurak M et al., 2007). The ability to detect differentially expressed genes suffers as a result.

Part of the Agilent pre-processing involves the use of a surrogate intensity value when the measured intensity of an array feature is not significantly different from background levels. This may contribute to the variability at the low threshold end.

Zahurak et al., (2007) also demonstrated that the largest portion of mean squared error for low intensity genes was due to random background estimation error rather than the systematic shrinkage of fold changes. It was found that this was true unless there was clear correlation between background and foreground, indicating that local background is an important component of foreground intensities.

In terms of the normalization process for microarray data, the purpose of normalization is to adjust for any bias which arises from variation in the microarray technology rather than from biological differences between the RNA samples or the printed probes. Most common is red-green bias due to differences between the labelling efficiencies and scanning properties of the two fluors complicated perhaps by the use of different scanner settings. Where only one colour of fluor is used this source of bias is irrelevant.

Other biases may arise from variation between spatial positions on a slide or between slides. Positions on a slide may vary because of differences between the print-tips on the array printer, variation over the course of the print-run or non-uniformity in the hybridization. Differences between arrays may arise from differences in print quality or from differences in ambient conditions when the plates were processed. It is necessary to normalize the intensities before any subsequent analysis is carried out. In our situation both within-array, between-array and between-slide normalization could be applied.

On the basis that we observed no systematic procedural bias (Figure 5.8), and in line with the accepted premise that normalization is intended to modify the underlying data

as slightly as possible in order to conserve biologically relevant information, but as much as necessary to remove systematic biases, the decision was made not to apply any further normalization as any potential benefit of such data processing could not be clearly identified a priori.

Whilst the rationale for not applying normalization in this instance has been clearly stated it should also be acknowledged that such an approach is not immune to criticism and that the subject of normalization, or at least the most appropriate normalization strategy for such one-channel arrays, remains controversial. Indeed, Fundel et al., (2008) illustrated that all normalization strategies make underlying assumptions of the data and their implementation can change the results of a particular experiment.

Another important issue with microarray studies is the requirement for some form of validation of the data. Ideally this would include a demonstration of coordinate changes at the protein level for gene(s) with significantly altered expression but it is also common practice to show similar changes in expression at the molecular level with a different assay system such as RT-PCR. In this study, antibody detection of Ki67 has already been used to illustrate proliferative changes. The microarray contained a probe for the proliferating cell nuclear antigen (PCNA), therefore it could be demonstrated that the molecular data for PCNA gene expression from the array showed significant correlation with the presence of Ki67 at the protein level.

The previous chapter also demonstrated significant changes in goblet cell number in the damaged and transitional areas. IL-4 and IL-13 were identified as genes present on the array that have been implicated in changes in goblet cell numbers and/or mucus production (Dabbagh K et al., 1999, Shim JJ et al., 2001). However no correlation between expression of these genes and GC numbers was observed. This is unlikely to be the sole or main function of these genes therefore it is by no means certain that such a relationship ought to be present or that changes would necessarily be mediated at the transcriptional level.

To further assess the reliability and reproducibility of the microarray data presented in this study, semi-quantitative RT-PCR assays were carried out on a selection of genes which demonstrated highly significant changes ($p<0.001$) in expression in injury sites. The selection of these genes were based on one of the following; 1) a distinct change in expression level at various time points as determined by high and low fold changes 2) different functional categories known to be modified during airway injury and repair, and 3) genes previously reported as critical mediators in airway inflammation. This study revealed that for all selected genes there is a significant correlation ($p<0.05$) between the array and the semi-quantitative RT-PCR data in terms of the expression profiles except the gene for beta-catenin ($p>0.05$). The significant correlation between array and semi-quantitative RT-PCR assay gives us confidence that the microarray data is reliable and reproducible even by using a non-probe based technique (semi-quantitative RT-PCR) for validation study.

The selected genes for validation study are thought to be involved in various functions in response to airway damage and repair. Interleukin 6 (IL6) was one of the most significantly up-regulated genes from the microarray analysis and can act either as a pro-inflammatory or anti-inflammatory cytokine. It is produced by the T cells and macrophages to stimulate an inflammatory immune response to tissue damage. IL6 expression was also reported to be down-regulated during airway epithelial cell apoptosis *in vitro* (Hodge S et al., 2002) and increased IL6 expression was important for mucus hypersecretion by airway epithelial cells in response to inhaled *Aspergillus fumigatus* (Neveu WA et al., 2009). The up-regulation of this gene from 6 hours to day 3 post injury observed in the microarray study was confirmed by the PCR data and indicates that this gene may play a role in the early response of the sheep airway to injury.

The KRT6A and KRT19 genes are known as markers of basal cells (Liu HW et al., 2009, Takahashi K and Coulombe PA, 1997, Wojcik SM et al., 2000) and an increase in

the number of poorly differentiated cells expressing basal cell markers has been associated with the early cellular response during the de-differentiation and migration processes (Evans MJ et al., 1989, Evans MJ and Plopper CG, 1988, Lane BP and Gordon R, 1974). The regulation of these genes might be consistent with a potential role for cells with basal cell-like properties as "progenitor" cells (Rock JR et al., 2009, Hong KU et al., 2004a, Voynow JA et al., 2005) during airway regeneration, particularly in migration and proliferation processes. Although there is a lack of studies describing the role of keratin 6A (KRT6A) and keratin 19 (KRT19) in airway repair following injury, these genes have been extensively studied in the injured skin (Takahashi K and Coulombe PA, 1997, Wojcik SM et al., 2000, Liu HW et al., 2009). It was suggested that the KRT6A was expressed by the basal cell and plays a role in activation of follicular keratinocytes after wounding (Wojcik SM et al., 2000). Mouse endogenous KRT6A is also reported to be expressed by the basal layer in epidermal tissue at the edge of the wound (Takahashi K and Coulombe PA, 1997). A study by Liu HW et al., (2009) using skin samples from aborted fetuses at different gestational ages and from healthy individuals, showed that the expression of KRT19 was mainly confined to the basal layer (Liu HW et al., 2009). Although at the time the primer was designed to amplify keratin 19, it was difficult to ensure that the designed primer was specific for the target gene especially in relation to the keratin genes. As clearly discussed in the Chapter 3 for the keratin 14 gene, majority of the keratin genes share highly similar mRNA sequence, therefore the chances that the designed primers could amplify other types of keratin genes are high. In addition to this, the limitation of the microarray and the semi-quantitative RT-PCR techniques may not have been able to distinguish between different keratins.

The up-regulation of these genes following injury in the microarray study was again confirmed by conventional PCR. This finding supports our initial hypothesis that basal-like cells are the first cells to repopulate the denuded areas following injury to form a single layer of squamous epithelial cells. It is possible that these have originated through dedifferentiation of epithelial cells bordering the lesion and early cellular migration

towards the centre of the lesion. The up-regulation of these genes with a peak at day 3 post injury suggest that these cells, with basal cell-like properties, play a major role during airway regeneration and repair. The histopathological changes of the airway at this time point showed that most of the lesion had repopulated with epithelial cells with evidence of some thickened areas of epithelial hyperplasia. We suggest that at this point in time, these basal-like cells may play a major role by giving rise to other types of epithelial cells (Hong KU et al., 2004a, Rock JR et al., 2009) (goblet and ciliated cells). The involvement of the basal cell in epithelial differentiation has been described previously (Boers JE et al., 1998, Evans MJ et al., 1989, Evans MJ and Plopper CG, 1988, Hong KU et al., 2004a, Voynow JA et al., 2005) as this cell was reported to play a role as stem or progenitor cell of the mouse trachea and human airway epithelium (Rock JR et al., 2009).

Other genes described in this validation study were genes involved in the cell cycle [cyclin B1 (CCNB1), cyclin B2 (CCNB2) and cell division cycle 2 (CDC2)] and the activation of genes involved in cell cycle was previously reported by Heguy et al. (Heguy A et al., 2007) in human airway epithelial repair following bronchial brushing. Transition from G_2 to mitosis is basically controlled by B-type cyclins in which cyclin B1 appears during S phase and is primarily cytoplasmic and moves into the nucleus just at the start of mitosis and associates with condensed chromosomes and the mitotic spindle (Pines J and Hunter T, 1992). In a study by (Heguy A et al., 2007) reported that most of the genes were highly expressed at day 7 post injury, predominantly at G_2 and M phases of the cell cycle phase (Heguy A et al., 2007). In the context of the present study, the activation of these cell cycle genes (CCNB1, CCNB2 and CDC2) from day 1 up to day 7 post injury with fold changes peaked at day 3, might suggest the active regulation of the cell cycle activity in response to repair process of the airway (Chapter 6_ Appendix 6.1 on CD) involving the transition from the S-phase to M-phase of the cell cycle activities. This finding is consistent with the Ki67 staining (Chapter 4) in which the proliferative activity was significantly increased at day 1 up to day 7 post injury as

compared to naive airway (time point 0) with various staining density of the Ki67-positive cells which indicate different stages of the cell cycle phase.

In regards to the beta-catenin gene (CTNNB), the expression profiles observed both in array and semi-quantitative analyses showed no significant correlation ($p>0.05$). However, consistent with our preliminary finding as described in Chapter 3, the expression of this gene detected by the semi-quantitative RT-PCR assay was showed to be up-regulated from the day 1 to day 7 post injury as compared to naive. This observation was in agreement with the expression of its known co-expressed gene, WNT5A which was significantly unchanged at day 3 whilst significantly increased 7 days post injury with the fold changes 1.9 and 2.1, respectively. Based on the array data, it appears that the expression of the CTNNBL1 gene was not affected following brushing.

As described in Chapter 4 for the 6 hour time point, the histopathological changes were characterised by disruption of the airway mucosal including the basement membrane. The single layer of squamous epithelial cells migrating from the border of the lesion is thought to be an early stage of the airway repair. At this time point, only few proliferative cells were observed as compared to other time points. At this time point the number of probes with significantly altered (both up- and down-regulated) expression compared to control was greater than other time points (4749 probes) at $p<0.05$ - (Table 5.8). However, it is noticeable that the majority of the altered expression was down-regulation whilst only 387 probes found to be up-regulated. An increased number of genes were significantly altered at day 3 compared to days 1 and 7 and was consistent with the histopathological finding in terms of proliferative activity and histopathological changes (epithelium hyperplasia/metaplasia and fibroplasias – (Chapter 2; Figure 2.9)) of the airway at this time point. This finding suggests that around day 3 may be a critical and important time for the airway to get to the stage where the repair machinery is being activated at both cellular and molecular levels. However, the proportion of up-regulated genes gradually increased over time following injury and by day 7, the numbers of up

and down regulated genes were relatively equal. Additional analysis to investigate functional classification and pathways for both up- and down-regulated genes will open a new window to better understand the biological function of these genes in regulating airway regeneration and repair.

In terms of the genes that have robust gene expression changes, our array data showed that the biggest fold change with up-regulation was of the MT1E gene (fold change of 110 at 6 hours) which gradually decreased down to 2.2-fold at day 7. This suggests that MT1E may play an important role in the early response to injury since it is so rapidly activated following injury. A study of endotoxin inhalation (Hur T et al., 1999) in the rat showed that the induction of this gene was greater (FC=4) after 3 hours exposure to the endotoxin and decreased to the basal level at 24 hour later. This gene was found to be largely increased in epithelial cells of the conducting airways, surrounding airway tissue, and in the nearby gas exchange region (Hur T et al., 1999). The function of this gene in the early response of the airway to bronchial brushing is still unknown, however it has been suggested that the metallothionein mRNA was inducted in the rodent liver (Andrews GK, 2000) to protect the cell from free radicals as the process of limiting free radical damage is a high priority for the cell survival. The major functions of the down-regulated genes with fold change more than 10 are involved in regulation of liquid surface tension, respiratory gaseous exchange, homeostasis process and defence response. Amongst those down-regulated genes, the expression of surfactant proteins (SP-A and SP-C) were significantly down-regulated from 6 hours up to day 3 post injury. Lung surfactant is a complex mixture of phospholipids and four surfactant-associated proteins (SP-A, SP-B, SP-C and SP-D). Its major function in the lung alveolus is to reduce surface tension at the air-water interface in the terminal airways by the formation of a surface-active film enriched in surfactant lipids, hence preventing cellular collapse during respiration (Walther FJ et al., 2007). Decreased levels of the main component of surfactant--phospholipids--have been implicated in atrophic rhinitis (Schlosser RJ, 2006). The lamellar body arrangement of phospholipids has now been demonstrated in both normal and diseased sinus tissue, resulting in the implication that

these structures may play a crucial role in mucociliary clearance against inhaled pathogens, as well as in the regulation of mucous viscosity (Schlosser RJ, 2006).

Since the advent of microarray technology it has become apparent that difficulties exist in reproducing results between, and often within, laboratories. Much of this difficulty may arise out of the inordinate number of non-biological variables that can influence results. Such variables range from the aforementioned technological issues surrounding printing, hybridisation, washing and scanning of slides to the various data processing steps that can be applied to correct for any inherent bias arising out of the technology. In recognition of this issue there has seen the inception of proposed standards that apply to the running and reporting of microarray experiments. In this regard Brazma et al., (2001) proposed such standards under the acronym "MIAME" in order to ensure that microarray data can be easily interpreted and that results derived from its analysis can be independently verified (Brazma A et al., 2001).

5.6 CONCLUSION OF THE CHAPTER

The microarray study has generated expression profiles for each of the time points following injury and provides information on which genes are up or down-regulated and by how much. The volume of data is such that it is impossible to comment on all the changes in expression observed and the challenge therefore is knowing how best to interpret and present it. Rather than trawling through these lists selecting genes either on the basis of the most significant fold-changes in expression or using a candidate-gene approach to select those which we think might be biologically relevant to the system under study, an alternative approach is to use the data to try and identify what pathways or networks might be involved through a process that applies state-of-the-art functional genomics and network analysis tools to link the experimental findings into functional gene networks with a much wider significance. The outcome of this alternative approach is described in the following chapter.

CHAPTER 6

ANALYSIS OF THE BIOLOGICAL FUNCTIONS OF SIGNIFICANTLY EXPRESSED GENES DURING SHEEP AIRWAY EPITHELIAL REGENERATION AND REPAIR

6.1 INTRODUCTION

Gene expression profiling through high throughput microarray analysis presents researchers with the considerable challenge of interpreting the biological relevance of any measured 'interesting' change in individual gene expression in the context of the system under study. Since the dawn of the genomic era the challenge of interpreting large lists of such interesting genes has been compounded by the masses of information that has found its way into many different database platforms. Whilst, in most instances, such information has been freely available – the manner in which such data was initially categorised in the process of recording created obstacles to researchers keen to assimilate and understand the information relating to specific biological processes. The lack of initial standardisation of terms used to describe gene products lay at the root of the problem – thus challenging attempts to assimilate data in a particular area where subtlety different terms could be used to describe essentially similar gene products. In recognition of this growing problem the Gene Ontology Consortium (founded in 1998) initiated the Gene Ontology (GO) project to seek the collaboration of databases in arriving at a consensus over the use of a shared vocabulary (GO terms) to be used to annotate gene products in the respective databases. Many tools have now been developed in order to facilitate the identification of relevant GO terms for lists of genes (Ashburner M et al., 2000). The Gene Ontology Consortium defines such GO terms using a controlled vocabulary that can be applied to all eukaryotic gene and protein functions. There are three main categories of GO terms – 1) the biological process (BP) - which refers to the biological objective to which the gene and gene product contributes, 2) the molecular function (MF) - which is defined as the biochemical activity of a gene product, and 3) the cellular component (CC) - which refers to the place in the cell that a gene product is active (Ashburner M et al., 2000).

By such unifying process it is possible to cluster genes based on their biological and biochemical functions. Further, it is possible to systematically dissect large 'interesting' gene lists in order to assemble a summary of the most enriched and pertinent biology.

The foundation of such enrichment is that if a biological process is more or less prominent in a given study, the co-functioning genes should have a greater potential for selection as a relevant group by the chosen screening technology. The significance of any enrichment can be assessed by common statistical methods, including Fisher's exact test, Chi-square, Binomial probability and Hypergeometric distribution. Moreover, the software tools required to facilitate this process have become increasingly available through open-access websites devoted to this aim. Current popular examples include GOstat (Beissbarth T and Speed TP, 2004), Fatigo (Al-Shahrour F et al., 2004), GenMAPP(Doniger SW et al., 2003) and DAVID (database for annotation, visualisation and integrated discovery) (Dennis G Jr et al., 2003). These programmes provide tools that can increasingly deal with requests originating from different species and can accommodate query inputs (the list of genes for which functional annotation is requested) formatted according to the identifiers used in the originating databases (including GenBank ID, Unigene, ENSEMBL, EntrezGene ID and GeneSymbols). These programmes also provide links that can direct the users to more specific pathway information (Beissbarth T and Speed TP, 2004, Huang da W et al., 2009, Huang da W et al., 2008), other GO terms that one single gene belongs to and also to the hierarchical pathway within which each specific given GO term exists. In the latter context it should be appreciated that the GO database is structured in a hierarchical form, where if a child term describes a gene product, then all of its (less specific) parent terms must also apply to that gene product. The level at which a gene is annotated in the GO hierarchy depends on the detail of knowledge about its biological behaviour the annotator had.

Traditional enrichment strategies take the users' genes of interest (selected on the basis of *p*-value and fold change) and then iteratively test the enrichment of each annotation term one-by-one in a linear mode. By predefinition of a *p*-value threshold the significantly enriched annotation terms can then be reported in a tabular format. Clearly the level at which the *p*-value threshold is set will have a major impact on the number of significantly enriched terms. As indicated by (Huang da W et al., 2008) the criteria governing selection of *p*-value thresholds is continually evolving and currently neither

any particular method nor opinion is widely accepted and appropriate for all situations. Rather, that in the context of enrichment analysis output, the results should be viewed in an advisory sense and that it is up to the user to interpret the results in the context of their biological 'sense'. Further potential error in interpretation can arise as a consequence of the effect of making so many individual tests at the same time. Standard statistical principles state that the more annotations tested, the greater the chance of an increase in the family-wide false-positive, or false-discovery rate (FDR). The FDR of a test is defined as the expected proportion of false positives among the given significant results (Benjamini and Hochberg, 1995). To control for such effects it is necessary to correct for such errors through application of methods such as Bonferroni or Benjamini–Hochberg (Shaffer, 1995).

In order to extend our ability to interpret the list of 'interesting' genes identified in Chapter 5 we sought to employ functional annotation and enrichment analysis techniques through application of the software tools that comprise the online DAVID resource (http://david.abcc.ncifcrf.gov/). The reasons for selecting this particular resource included the following: 1) DAVID is a freely available web-based programme, 2) DAVID provides user friendly tools that ease access to all relevant information relating to the terms that associate with the single genes, and provides the functional hierarchy within which genes are placed, 3) DAVID's methods are applicable for smaller gene lists, and 4) DAVID provides for different species of interest to be used as reference backgrounds. However, the analysis of biological function using DAVID programme is only a theoretical analysis of what the DAVID software is capable of and that is demonstrate the way in which microarray data can be interpreted. The valid conclusion of this theoretical analysis can only be done on the laboratory test. However, it is to note that the dataset used to understand the biological function and pathways analyses was from the gMeanSignal (raw data) files obtained from the microarray experiments of Chapter 5 to understand the process of pathway biology.

6.2 OBJECTIVES

The aims of this chapter were:

1. To cluster the significantly expressed genes detected by the microarray experiment undertaken in Chapter 5 based on their biological function using DAVID (http://david.abcc.ncifcrf.gov/)
2. To analyse the biological functions of the significantly expressed genes based on their GO terms relating to biological process, cellular component and molecular function and relate these functions to the staging of the repair process.
3. To determine the pathways associated with the list of significantly expressed genes

6.3 METHODS

6.3.1 Genes selection

Thirteen thousand four hundred and ninety-five probes were significantly regulated (either up or down) at $p<0.05$. Amongst those, 6,420 probes were annotated and these probes were used for downstream application of biological function and pathway analyses. The appendix (Chapter 6_ Appendix 6.1_Significant at p less than 0.05_annotated on CD) contains the list of significantly expressed probes (annotated) at $p<0.05$.

6.3.2 Curating ontologies through cross-species analogy

In order to identify human refseq RNA sequences highly homologous with the probe targets on the Agilent 15K ovine array a strategy of using a locally installed Basic Local Alignment Search Tool (BLAST) client was adopted which facilitated the submission of

batch queries to the NCBI BLAST servers that power the NCBI web BLAST service (www.ncbi.nlm.nih.gov/BLAST/).

The client, named blastcl3, bypasses the web browser and is capable of performing the batch search with multiple sequences by taking one query sequence at a time from the input file, formulating the search according to the settings of the command line parameters, and sending the search through the internet connection to NCBI BLAST server for processing. The program receives the search results from the BLAST server, in the format set in the search command line, and saves it to a local file specified. The program loops through all the queries in the input file until all are searched.

Because the only query format blastcl3 accepts is FASTA it was necessary in the first instance to use entrez batch (http://www.ncbi.nlm.nih.gov/sites/) to upload, in batches, the list of accession numbers and output the respective FASTA sequences to text files – with a separate file for each consecutive 500 sequences (FASTA 1_500.txt, FASTA 501_1000.txt, FASTA 1001_1500.txt, et seq.).

Because the blastcl3 program has no graphic user interface (GUI) and must be executed from command line under a terminal window it was necessary to programmatically create a command line executable that would automatically read a specified number of FASTA entries at a time from each file and submit these as a query to NCBI BLAST using the blastcl3 client. The program, “DDSC Blastcl3 recaller.vi” kindly made available by Dr D Collie, was created to meet these needs using LabView (National Instruments, USA). The specific command line submitted at each call was as follows:

blastcl3 -p blastn -i FASTA_input.txt -p blastn -d refseq_rna -o n_refm.txt -e 2e-20 -m 8 -I T -u "human[organism]"

The various parameters included in the command reflect the options available to define the nature of the search and are explained as follows:

blastcl3	The name of the client .exe file located on the host computer
-p	A mandatory default parameter which simply specifies which program to run – in this case – the following "Blastn"
Blastn	Specifies that the Blastn search algorithm is to be used – ie a nucleotide blast on nucleotide databases
-I	Specifies that the accession number, GI, should be shown in the definition line
FASTA_input.txt	The name of the text file containing the FASTA sequences for BLAST submission
-d	Specifies the target database(s) to search against – in this case – the following "Refseq_rna"
Refseq_rna	The target database to be searched – in this case the Reference Sequence RNA collection.
-o	Specifies the output file – in this case – the following "N_refm.txt"
N_refm.txt	The output text file
-e	Specifies the Expect value cutoff
2e-20	The expect value used in this instance ie $2x10^{-20}$
-m 8	Specifies the alignment view option – in this case in a tabular format
T	Specifies whether to produce HTML output (defaults to false)
-u	Restricts the search of the database to the subset satisfying the query
"human[organism]"	Specifies the human subset in the refseq_rna database

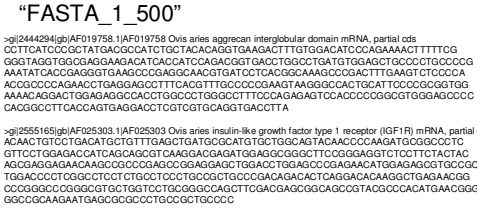
For each file of FASTA sequences a corresponding file listing the identities of entries in the RefSeq_RNA database that met the search criteria was created (FASTA 1_500_hits.txt, FASTA 501_1000_hits.txt, FASTA 1001_1500_hits.txt, et seq.). These separate files were collated into one file "Collated FASTA hits.txt" containing over 30,000 entries and these entries entered in a database file within Microsoft Access (Microsoft Office, Microsoft, UK).

It was then possible to reduce the burden of non-essential entries by creating a query within Microsoft Access to pick out, for each unique query sequence identity, the FASTA hit with the lowest expect value where the homologous refseq entry is prefixed by "NM" – for manually curated information. This created a list comprising 8901 entries. Limiting the query to select only the manually curated information for mature mRNA transcripts filtered out reference to non-coding transcripts (NR) and coding and non-coding transcripts (XM and XR) identified on the basis of automated procedures.

The NM identifier could then be used as input into IDconverter, the online tool available for conversion and annotation of gene and protein IDs (http://idconverter.bioinfo.cnio.es/), to yield the Gene Name (taken from UniGene), Ensembl transcript ID, UniGene ID, EntrezGene ID, the relevant GenBank Accession numbers and Clone IDs and information relating to chromosomal position.

This information could then be paired with the originating query sequence identities to provide the means by which the ascribed annotations could be linked to the results of the array hybridizations and, following addition of any manually curated annotations, thereafter made available for forward analysis. The process is depicted in the figure below:

"FASTA_1_500"



Blastcl3

Blastcl3 recaller

ID	Query Sequence Identity	Matching sequence identity	Percent Identity	Alignment length	# mismatches	# gaps	Start of query sequence	End of query sequence	Start of matching sequence	End of matching sequence	E value	Bit score
1	gi|47950003|gb|CN821934.1|CN821934	gi|194097329|ref|NM_002115.2|	87.78	581	71	0	128	708	83	663	3E-167	589

"FASTA_1_500_hits"
↓
Collated FASTA hits

Best matching seqs_Collated FASTA (20090414) Query

Query Sequence Identity	Query Sequence ID	MinOfE value	CountOfE value	FirstOfMatching sequence identity	Extract1	NM Identifier
gi|10086316|gb|AF283510.1|AF283510	AF283510	6E-117	21	gi|212549565|ref|NM_171825.2|	NM_171825.2|	NM_171825
gi|10086318|gb|AF283892.1|AF283892	AF283892	1E-73	3	gi|117320539|ref|NM_001077493.1|	NM_001077493.1|	NM_001077493

http://idconverter.bioinfo.cnio.es/IDconverter.php

FASTA hit identities (20090414) table

ID	Gene Name	Ensembl Gene	UniGene	EntrezGene	RefseqRNA	GenBank Accession	Clone ID	Ensembl* Chr	Start (bp)	End (bp)	Strand
2	GSS	ENSG00000100983	Hs.82327	2937	NM_000178	BC007927 NM_000178 CR602068 U34683 CR596201 CR617218 CR620374 CR611077 BQ952755 BI757536 BF690367 BM557033 BQ649135 BQ940490 BQ897884 BE882644 BU507992 BQ067914 BU539684 AL563039 BE561366 BG033543 BX375152 BQ931334 AL553227 BQ953111 BQ071237 BQ051368 BQ06	IMAGE:6484315 IMAGE:5199670 IMAGE:4299112 IMAGE:5466525 IMAGE:6286187 IMAGE:6338418 IMAGE:6084319 IMAGE:3912601 IMAGE:6502016 IMAGE:5768893 IMAGE:6570487 CS0DC028YG09 IMAGE:3677816 IMAGE:4403796 CS0DC009YG09 IMAGE:6484398 CS0DI073YN07 IMAGE:6375928 IMAGE:	20	32979898	33007262	-1

Pair best matching seq with FASTA hit identity info: Query

Human homolog	Query Sequence ID	MinOfE value	Count OfE value	Matching sequence identity	NM Identifier	Refseq RNA	Ensembl Gene	UniGene	EntrezGene	GenBank Accession	Clone ID	Ensembl * Chr	Start (bp)	End (bp)	Strand
FGD2	DY509924	2E-82	1	gi|189217916|ref|NM_173558.3|	NM_173558	NM_173558	ENSG00000146192	Hs.509664	221472	AK131079 AK024456 NM_173558 BC023645 BX648164 AK097981 AK092732 AK097230 AK290419 BC053655 AK098248 BC062363 BM922781 BQ705924 BQ708371 BG398458 BM007617 AI524250 CD366726 BG686784 DR004735 EL734772 DR006935 AI524728 DA619300 DA943324 DA337275 CA392949 DA	IMAGE:5755893 IMAGE:6214791 IMAGE:6215310 IMAGE:4565722 IMAGE:5440975 IMAGE:2117996 UI-H-FT2-bjo-p-13-0-UI IMAGE:4763317 TC104804 TC105056 IMAGE:2118091 JCMLC2000144 SPLEN2018181 BRHIP3038141 cs30h10 PEBLM2006774 SMINT2014620 D9OST2002477 BRAMY2046111 SPL	6	37081400	37489422	1

Results aveffect plus all annotation: Query

Combined annotation	Human homolog	Probe names	ProbeID	TargetID	Gene Symbol	Gene Name	Accessions	Description	aveffect_1
ADA		A_70_P000001	9080	EE850561	ADA	adenosine deaminase	ug|EE850561	020829OSI1005015HT OSI Ovis aries cDNA, mRNA sequence [EE850561]	0.314693031
		A_70_P000011	8399	EE751926			ug|EE751926	020314OCSI1001069HT OCS1 Ovis aries cDNA, mRNA sequence [EE751926]	-0.725248451
ZW10	ZW10	A_70_P015006	893	DY494304			ug|DY494304	sh2P0012F21_F.ab1 adult sheep fracture callus 7d Ovis aries cDNA, mRNA sequence [DY494304]	-0.192484002

etc

Figure 6.1: Flow chart representing the process involved in curating ontologies through cross-species database interrogation.

6.3.3 Identification of GO terms and gene functions of significantly expressed genes

Gene functional enrichment analysis was performed using a free web-based programme, DAVID version 2008 (http://david.abcc.ncifcrf.gov/), as previously described (Dennis G Jr et al., 2003, Huang da W et al., 2009, Huang da W et al., 2007, Huang da W et al., 2008). Two lists of genes are specified in setting up the analysis using this programme, the first being the genes of interest represented on the microarray, and the second being a background set of genes. The latter is required in order to assign the associated GO terms relating to the list of significant genes. The complete GO gene-associations database for the human genome is used as a reference in this regard. Whilst the preferred background is the entire genome-wide genes of the species concerned, and prebuilt backgrounds, such as the genes represented on Affymetrix chips, are available for the user's choice (particularly for well-studied model organisms such as human, mouse, rat and *Arabidopsis*), the relatively poor annotation of the Agilent sheep array conspired against using sheep genes as the background of choice in this study. A presumption was therefore made that human biological pathways are sufficiently similar to those of the sheep and effort was directed towards identifying the nearest human homologs to the significantly expressed genes in order that the human genomic database could be used as the population background. Selecting a larger set of background genes tends to give smaller *p*-values for each set of the given functional classification (Huang da W et al., 2009). In the analysis, DAVID determines all the annotated GO terms that exist both in the background gene list and those associated with our genes of interest. The number of appearances of each GO term will be counted and compared between the groups of interest and for the reference genes. A modified Fisher exact test was calculated for all analyses and the Benjamini-Hochberg method (Benjamini and Hochberg, 1995) was used to control the false discovery rate for the enrichment *p*-values for the given individual term members.

Two approaches were adopted in order to analyse our lists of genes. The first ‘general analysis’ involved screening for significant GO terms associated with our total set of significantly expressed probes (6,420 annotated probes). The second

'specific analysis' was performed in two ways, the first involved considering the significantly regulated genes at each time point separately, and the second involved separating the significantly regulated genes at each time point into lists according to whether they were up-, or down-regulated. For the latter analysis, largely prompted by the disproportionate distribution of down-regulated genes at the 6 hour time point (Chapter 5; Figure 5.12), up- and down-regulated genes were considered separately and inference gained as to the major processes involved as a consequence. However, the "general analysis" was done only to get overview of the general GO terms involved in our list of significantly annotated genes whilst "specific analysis" for both types of genes regulation (up- and/or down-regulated) was used to get more specific and relevant biological functions of the annotated probes when probes were separated by the time points especially for the pathway analysis.

An important consequence of the latter approach was a reduction in the number of genes presented for forward analysis, and whilst such gene lists could still be clustered according to function, the associated significance of such groupings suffered some erosion and this was generally reflected in increased *p*-values for each given term.

Using this tool, the relationship between the given lists of the annotation terms were measured on the basis of their co-association with the genes expressed at the particular time points and clustered based on their similar function. DAVID default settings were used. Such default settings are recognised to increase the specificity and tightness of the analysis and reduce the overall number of clusters (Huang da W et al., 2009). As an initial guideline for selecting an appropriate degree of clustering to identify time-dependent shifts in biological process, molecular function and cellular components during repair, the significance of gene-term probabilities was set at $p<0.05$. However, attention was also directed towards the appraisal of non-significantly enriched terms in order to avoid overlooking potentially important biology.

6.3.4 Pathway analysis

The DAVID programme was also used to consider the relevant pathways that could be involved in the given sets of genes. Particular interest in this regard was directed towards identifying and graphically summarizing the distribution of the associated genes among functional domains and pathways in KEGG metabolic and regulatory pathways (Dennis G Jr et al., 2003). Similar to the strategy adopted in identifying the associated GO terms (Section 6.3.3 in this chapter), an initial "general analysis" was undertaken, followed by time point of "specific analyses". As previously noted the strength of relationship between likely pathways and gene lists was influenced by the numbers of gene involved, with larger gene sets producing smaller significant probabilities (*p*-values).

6.4 RESULTS

6.4.1 Analysis of the biological functions of differentially expressed genes

Initial analysis was performed by combining all significantly expressed genes from all the time points. This analysis was performed in order to construct an overview of 1) the major biological functions associated with the gene list, 2) to discover the enriched functionally related genes and 3) to explore genes-terms relationship in the list of significantly expressed genes. In order to make the resulting group specific GO terms more interpretable the results were clustered. Following identification and association of available homologous human gene symbols to the combined 6420 significantly regulated probe targets and input of this information to the DAVID functional analysis tool, 2550 unique genes could be identified - the reduction in number reflecting both probe redundancy and a failure of the DAVID human database programme to recognise the annotated gene symbols. In the latter context it should be appreciated that in some instances, the given gene annotations were only available as the locus names e.g. LOC514296, or species-specific forms, such as "C11H9orf16" for the manually curated bovine chromosome 9 open reading frame 16 ortholog gene, which were not recognised by the DAVID programme. The advantage of using DAVID is that the programme converts the given gene symbols to a unique format recognised by the database and these unique gene symbols used

for downstream analysis using DAVID. The list of genes converted by the DAVID programme is shown in the (Chapter 6_ Appendix 6.2_DAVID conversion on CD). This set of genes was clustered according to their respective GO terms. The lists of significant terms (BP, CC and MF) are shown in the appendix (Chapter 6_ Appendix 6.3(a) to (c) on CD).

The most significant functions represented in the gene list were those relating to cellular component organization and biogenesis, positive regulation of biological process, response to stress, developmental process and cellular process. Other highly represented functions included cell proliferation, cell cycle, cell differentiation, mitotic cell cycle, inflammatory response, blood vessel development, response to wounding, vasculature development, apoptosis and cell development. Some terms such as cell activation, transcription from RNA polymerase II promoter, cell migration, regulation of B cell activation, lymphocyte differentiation, cell-matrix adhesion, cytokine metabolic process, cytokine production, embryonic development, T cell proliferation and regulation of cell differentiation were amongst the less significant clusters.

The list of pathways and component gene counts are shown in the (Table 6.1) and the list of genes that fall in each pathway are shown in the appendix (Chapter 6_ Appendix 6.3(d) on CD). Twelve pathways identified by DAVID commonly associate with our gene list – the most significant being those involved in cell cycle and ECM-receptor interaction. The genes were grouped in these pathways based on shared or similar functions/terms.

Term	Count	%	*p*-value	Benjamini
Cell cycle	46	1.8	6.40E-07	1.30E-04
ECM-receptor interaction	38	1.5	1.10E-06	1.10E-04
Focal adhesion	63	2.5	5.70E-05	3.80E-03
Small cell lung cancer	30	1.2	1.80E-03	8.50E-02
Leukocyte transendothelial migration	37	1.4	2.60E-03	9.90E-02
Cell communication	40	1.6	6.80E-03	2.10E-01
p53 signalling pathway	23	0.9	9.00E-03	2.30E-01
TGF-beta signalling pathway	28	1.1	1.20E-02	2.70E-01
Hematopoietic cell lineage	26	1	2.00E-02	3.60E-01
Toll-like receptor signalling pathway	28	1.1	5.70E-02	6.60E-01
Complement and coagulation cascades	20	0.8	7.20E-02	7.10E-01
Cytokine-cytokine receptor interaction	61	2.4	7.40E-02	7.00E-01

Table 6.1: The list of pathways associated with the complete list of significantly regulated list of genes (combined analysis).
The most highly significantly associated pathways are listed at the top of the list. Although the *p*-values for the toll-like receptor signalling, complement and coagulation cascades and cytokine-cytokine receptor interaction pathways are not significant, these pathways are considered biologically relevant to airway regeneration and repair, and thus taken for downstream application throughout the analysis.

Keys:

Term:
Relevance KEGG pathways associated to the given genes analysed using DAVID database. The most significant terms are listed on top of the list.

Count:
The number of genes clustered in the given pathways

%:
The percentage of the genes clustered in the given terms relative to the number of analysed genes i.e 46 genes clustered in the cell cycle pathways / 2550 of total unique genes analysed X 100 = 1.8%.

P-value:
Modified Fisher Exact *p*-value, EASE score - calculated by DAVID. DAVID determined all pathways associated with the list of genes analyzed. It will then count the number of appearances of each pathway for the genes inside the group and for the reference genes (human genomic database). Fisher's Exact Test is performed to judge whether the observed difference is significant or not. The smaller, the more enriched.

Benjamini:
A false discovery rate; the values presented have been multiplied by 100 (range from 0 to 100%). Therefore, a value of 0.1 represents one tenth of one percent.

The analysis was repeated for each time point separately.

6.4.1.1 GO terms

6 hours

One thousand, six hundreds and sixteen unique genes were significantly differentially expressed at this time point. Among these, 1286 genes (79%) were clustered in the GO term of biological process, 1339 genes (82%) were clustered in the GO term of cellular component and 1370 genes (84%) were clustered in the GO term of molecular function. The lists of significant terms associated with genes expressed at 6 hours time point are shown in the appendix with the most significant GO terms are listed on top of the list [(Chapter 6_ Appendix 6.4a_Biological process_6 hours), (Chapter 6_ Appendix 6.5a_Molecular function_6 hours), and (Chapter 6_ Appendix 6.6a_Cellular component_6 hours) on CD]. The most significant GO terms for the biological process are positive regulation of biological process, positive regulation of cellular process, developmental process, cellular component organization and biogenesis and response to stress. A further analysis was completed to classify the functions of the genes separated based on whether they were significantly up- or down-regulated after injury. Separation according to these criteria established that 281 unique genes were up-regulated whilst 1338 unique genes were down-regulated at the 6-hour time point. Amongst those up-regulated were genes significantly involved in response to stress, response to wounding, response to external stimulus, inflammatory response and amine metabolic process [(Chapter 6_ Appendix 6.7a) on CD]. The down-regulated genes[(Chapter 6_ Appendix 6.7b) on CD] at this time point were most significantly involved in positive regulation of biological process, positive regulation of cellular process, developmental process, biological regulation and cellular component organization and biogenesis (Table 6.2).

Day 1

One thousand and ninety eight unique genes were significantly expressed at this time point with 919 genes (83%) clustered in the GO term of biological process, 949 genes (86%) clustered in the GO term of cellular component and 949 genes (86%)

clustered in the GO term of molecular function. The lists of significant terms associated with genes expressed at day 1 time point are shown in the appendix with the most significant GO terms listed at the top of the list [(Chapter 6_ Appendix 6.4b_Biological process_ Day 1), (Chapter 6_ Appendix 6.5b_Molecular function_ Day 1), and (Chapter 6_ Appendix 6.6b_Cellular component_ Day 1) on CD]. The most significant GO terms for biological process included cell cycle process, cell cycle, response to stress, mitotic cell cycle and regulation of progression through cell cycle. Implementing a further analysis with gene lists separated according to whether significantly up- (642 genes) or down- (459 genes) regulated indicated that the most significant GO terms for the biological process for up-regulated genes [(Chapter 6_ Appendix 6.8a) on CD] were cell cycle, cell cycle process, mitotic cell cycle, mitosis and M phase of mitotic cell cycle whilst the most significant GO terms for the biological process of down-regulated genes [(Chapter 6_ Appendix 6.8b) on CD] were those concerned with developmental process, anatomical structure development, cell differentiation, cellular developmental process and multi-cellular organismal development (Table 6.2).

Day 3

One thousand, two hundred and forty nine unique genes were significantly expressed at this time point with 1030 (82%) clustered in the GO term of biological process, 1068 (85%) clustered in cellular component and 1050 (84%) clustered within molecular function. The lists of significant terms associated with genes expressed at day 3 time point are shown in the appendix with the most significant GO terms listed at the top of the list [(Chapter 6_ Appendix 6.4c_Biological process_ Day 3), (Chapter 6_ Appendix 6.5c_Molecular function_ Day 3), and (Chapter 6_ Appendix 6.6c_Cellular component_ Day 3) on CD]. The most significant GO terms for biological process associated with the list of significant genes at this time point were response to stress, mitosis, mitotic cell cycle, M phase of mitotic cell cycle and M phase. At this time point, there were 646 up-regulated genes whilst 610 genes were down-regulated. In considering these gene lists separately, the most significant GO terms amongst the up-regulated genes [(Chapter 6_ Appendix 6.9a) on CD] were mitotic cell cycle, mitosis, M phase of mitotic cell cycle, M phase and cell cycle

phase, whilst the most significant GO terms amongst the down-regulated genes [(Chapter 6_ Appendix 6.9b) on CD] were biological regulation, cellular process, response to stress, regulation of biological process and developmental process (Table 6.2).

Day 7

Eight hundred and thirty eight unique genes were significantly expressed at this time point with 691 (82%) clustered in the GO term of biological process, 700 (83%) clustered in cellular component and 713 (85%) clustered in molecular function. The lists of significant terms associated with genes expressed at day 7 are shown in the appendix with the most significant GO terms listed at the top of the list [(Chapter 6_ Appendix 6.4d_Biological process_ Day 7), (Chapter 6_ Appendix 6.5d_Molecular function_ Day 7), and (Chapter 6_ Appendix 6.6d_Cellular component_ Day 7) on CD]. The most popular GO terms for biological process were mitosis, M phase of mitotic cell cycle, developmental process, organ development and M phase. Amongst those genes significantly expressed at this time point were 372 up-regulated genes and 470 down-regulated genes. The most significant GO term for biological process in up-regulated genes (Chapter 6_ Appendix 6.10a) were those involved in M phase, mitosis, M phase of mitotic cell cycle, mitotic cell cycle and phosphate transport while those for down-regulated genes (Chapter 6_ Appendix 6.10b) were involved in metabolic process, cellular metabolic process, muscle system process, muscle contraction and cellular biosynthetic process (Table 6.2).

Up-regulated genes

Time points	GO terms		
	Biological process	**Cellular component**	**Molecular function**
6 hours	response to stress	extracellular region	protein binding
day 1	cell cycle	organelle part	protein binding
day 3	mitotic cell cycle	extracellular region part	protein binding
day 7	M phase	extracellular region part	extracellular matrix structural constituent

Down-regulated genes

Time points	GO terms		
	Biological process	**Cellular component**	**Molecular function**
6 hours	positive regulation of biological process	Cytoplasm	protein binding
day 1	developmental process	integral to plasma membrane	protein binding
day 3	biological regulation	Cytoplasm	protein binding
day 7	metabolic process	intracellular part	protein binding

Table 6.2: Lists of the most significant GO terms for biological process, molecular function and cellular component associated with the up- or down-regulated genes at each specific time point.

6.4.1.2 Pathways

6 hours

Five hundred and twenty-two (32%) out of the 1616 significantly expressed unique genes were associated with KEGG pathways. The most significantly highly associated pathways involved at this time point were small cell lung cancer and TGF-beta signalling pathways. The details for the pathways associated with the genes regulated at 6-hour time point and genes clustered in each pathway are shown (Chapter 6_ Appendix 6.11(a)_Pathways_ 6 hours on CD).

Day 1

Three hundred and ninety-six (36%) out of the 1098 unique genes at this time point were associated with KEGG pathways. The most significantly highly associated pathways involved at this time point were the cell cycle and p53 signalling pathways. The details for the pathways associated with the genes regulated at day 1 time point and genes clustered in each pathway are shown (Chapter 6_ Appendix 6.11(b) _Pathways_ Day 1 on CD).

Day 3

At this time point, 421 (33%) out of the 1249 unique genes were shown to be associated with KEGG pathways. The ECM-receptor interaction and focal adhesion pathways were amongst the most significantly highly associated pathways involved at this time point. The details for the pathways associated with the genes regulated at day 3 time point and genes clustered in each pathway are shown (Chapter 6_ Appendix 6.11(c) _Pathways_ Day 3 on CD).

Day 7

Three hundred and seven (36%) of the 838 unique genes were associated with KEGG pathways. The ECM-receptor interaction and cell communication pathways were the most significantly associated pathways involved at this time point. The details for the pathways associated with the genes regulated at day 7 time point and genes clustered in each pathway are shown (Chapter 6_ Appendix 6.11(d) _Pathways_ Day 7 on CD).

A summary of the pathways involved in each time points is shown in the (Table 6.3).

(Chapter 6_Appendix 6.12(a-k) – as attached) shows a schematic of the relationship between the pathways that are associated with the significant genes lists (combined time points).

	TIMEPOINTS			
	6 hours	**Day 1**	**Day 3**	**Day 7**
PATHWAYS	Small cell lung cancer	Cell cycle	ECM-receptor interaction	ECM-receptor interaction
	TGF-beta signalling pathway	p53-signalling	Focal adhesion	Cell communication
	Hematopoietic cell lineage	Complement & coagulation cascade	Cell cycle	Focal adhesion
	Cell cycle	Small cell lung cancer	Cell communication	Cell cycle
	Leukocyte transendothelial migration	ECM-receptor interaction	Complement & coagulation cascade	Complement & coagulation cascade
	ECM-receptor interaction		p53-signalling	Leukocyte transendothelial migration
	Toll-like receptor signalling		Toll-like receptor signalling	Adherens junction
			Small cell lung cancer	Cell-adhesion molecules (CAMs)
				Hematopoietic cell lineage
				p53-signalling

Table 6.3: A summary of the time-dependent changes in pathways associated with the repair process
The most significant pathways are shown on top of the list.

6.5 DISCUSSION

Conventional microarray data analysis often yields comprehensive lists of genes that are associated with experimental intervention in some way. The burden that is placed on scientists is the interpretation of what, amongst this unwieldy listing, is actually biologically significant. As biological effect normally relates to the coordinate interplay of multiple genes acting along complex pathways it pushes the scientist to consider microarray data in the context of both the roles that the gene products are perceived to have in biology, and also in relation to how such underlying biology can be pieced together from studying the coordinate expression of genes of similar function. This chapter is constructed with this in mind.

One important issue that deserves some discussion concerns the selection of the background for the functional analysis of the gene expression data. Based on the current, relatively poor annotation of the sheep genomic database, the classification of sheep gene functions was determined by analogy with similar genes characterised in the human genomic database. The assumption would be that underlying mechanisms are indeed relatively consistent between these species. As previously noted (Chapter 1; Section 1.10) sheep are generally well regarded as animal models for human disease – reflecting the similarities in structure, function and biology in health and disease. As an example in relation to the lung, a combined burn and smoke inhalation injury in sheep model has been reported to mimic human acute respiratory distress syndrome (ARDS) secondary to multiple trauma and airway damage (Westphal M et al., 2006) and the sheep is also considered an ideal potential model for cystic fibrosis (CF) study (Broackes-Carter FC et al., 2002 , Tebbutt SJ et al., 1996, Tebbutt SJ et al., 1995). The sheep lung has been well studied from both electrophysiological and developmental aspects and shows close similarities to the human lung, particularly with respect to the function of the respiratory epithelium (Davies LA et al., 2008, Olver RE and Robinson EJ, 1986). It has also been shown that spatial and developmental expression of CFTR in the sheep is similar to that seen in the human and thus the sheep lung may provide a useful system in which to reproduce CF lung disease (Tebbutt SJ et al., 1995). In this context sheep are currently being used in preclinical testing of gene therapy vectors (Davies LA et al., 2008).

In the present study, a set of significantly (p<0.05) altered annotated probes were clustered based on their shared and similar GO terms. The use of gene symbols for downstream DAVID analysis imposed limits in that some genes were assigned synonyms or aliases which were not recognised by the programme or for other genes only locus positions were available. Amongst 2550 unique probes/genes, only around 2389 genes matched with human database which recognised by DAVID were used.

DAVID was first introduced by (Dennis G Jr et al., 2003). This programme translates the list of significantly expressed genes into a functional profile which provides insight into the cellular mechanisms relevant, in the context of the present study, to airway regeneration and repair. Many tools are now available to undertake functional analysis of microarray gene expression data. These include GeneSpring (Agilent Technologies Inc), GoStat (Beissbarth T and Speed TP, 2004), Fatigo (Al-Shahrour F et al., 2004), Blast2GO (Conesa A et al., 2005), GenMAPP (Doniger SW et al., 2003), KEGG (http://www.genome.jp/kegg/) and Ingenuity (http://www.ingenuity.com). Most of the tools are freely available web-based programmes and some require a licence (GeneSpring and Ingenuity software). Although these tools use the same core strategy to systematically map a large number of interesting genes to the list of associated biological annotation, they differ greatly in many aspects that influence both the utility and outcome of the analysis (Khatri P and Drăghici S, 2005). In the former regard the degree to which tools offer utilities ranging from pathway analysis to ID conversion varies and this can influence the nature and extent of analysis undertaken. In the latter regard it is worthwhile considering the different methodologies that underlay the core strategy.

Huang da W et al., (2008) describes three major classes of enrichment algorithms: 1) singular enrichment analysis (SEA), 2) gene set enrichment analysis (GSEA), and 3) modular enrichment analysis (MEA). In the SEA class, the enrichment p-value is calculated on each term from the pre-selected interesting gene list (e.g. differentially expressed genes selected between experimental versus control samples which were filtered based on p-value and fold changes). This strategy is commonly employed, reflecting both its simplicity in technique and inherent efficiency. However, the

linear output of terms can be very large and overwhelming, potentially diluting focus and obscuring relevant interrelationships (Huang da W et al., 2008). The GSEA approach involves a 'no cut-off' strategy and uses expression data relating to all the genes from a microarray experiment – regardless of any associated significance. The benefits of using this technique are that it reduces the arbitrary factors involved in traditional enrichment analysis and ensures that all genes contribute by taking every gene-specific experimental value (e.g. fold change) into the calculation for each annotation term (Huang da W et al., 2008). The 'no cut-off' strategy is the key advantage of GSEA, but is also becoming its major limitation in many biological studies. Indeed whilst the GSEA method requires a summarized biological value (e.g. fold change) for each of the genome-wide genes as an input, the relevance (whether statistical or biological) of any given fold-change is unknown. The technique relies on mapping patterns of association between groups of genes behaving in a similar way (Huang da W et al., 2008). The MEA strategy inherits the basic enrichment calculation found in SEA and incorporates extra network discovery algorithms by considering the term-to-term relationships in which joint terms may contain unique biological meaning for a given study and building such relationships is one step closer to the true nature of biology during data mining. Similar to the SEA approach, the use of a pre-selected gene list impacts on the analytic results (Huang da W et al., 2008).

In the context of the present study, we used the DAVID programme as it is based on the SEA and MEA strategies. The relatively poor annotation of the sheep array was a factor in not considering the GSEA technique a realistic option in the current study. DAVID takes into account the redundant and networked nature of biological annotation content in order to concentrate on building the larger biological picture rather than focusing on individual terms or genes – and as a consequence moves us a step closer to understanding the nature of biological process (Huang da W et al., 2008).

In relation to the process of identifying which programmes to use for functional analysis of the data set generated as part of the present study both GOstat and

DAVID programmes were considered user friendly and both offered a minimum limit on the number of genes to be considered. Another particular advantage of these two programmes pertains to their accommodation of analyses involving relatively few annotated probes. However one perceived limitation of using GOstat in this context was that this programme would only provide lists of significant GO terms and their associated genes but did not provide the necessary tools to embark on further analysis e.g into pathway representation. In addition, DAVID provides some unique features and capabilities, such as an integrated and expanded back-end annotation database – the DAVID Knowledgebase - involving a single- linkage method to agglomerate tens of millions of gene/protein identifiers and associated annotation from dozens of well-known bio-databases, advanced modular enrichment algorithms and powerful exploratory ability in an integrated data-mining environment (Huang da W et al., 2009).

Regardless of the specific methodology employed in the analysis, data interpretation is ultimately critical in the sense that the data generated merely reflects the input to the various algorithms employed, and making sense of the output must rely on biological plausibility and hence demands grounding in biological science. Indeed, although the DAVID programme provides significant *p*-values for given GO terms and FDR in order to increase confidence in output, these values are just guidelines and should not be used as the sole arbiter of the decision-making process (Dennis G Jr et al., 2003, Huang da W et al., 2008). Although some of the suggested pathways were not shown to have significant *p-v*alues their ranked importance appears to be biologically plausible when considered in the context of the associated phenotypic changes underway in the brushed airway wall.

In the present study, the classification and analysis of the genes based on their regulation following injury (either up- or down-regulated) at each time point provided valuable information in relation to how genes and pathways inter-connected with each other during airway injury and repair. The biological functions of the genes which were up-regulated early after injury were those involved in the early cellular adaptation to the injury such as response to stress and wounding and also

activation of innate immune mechanisms (6 hours). Later on the most significant functions were those associated with cell cycle regulation. These findings are in the agreement with the concomitant histopathological changes highlighted in Chapter 4. Whilst some of the up-regulated genes were involved in cell cycle activities and mitosis and those genes defined as down-regulated were mostly involved in control of airway homeostasis which associate with developmental process, anatomical structure and development, metabolic process, cellular regulation and cellular differentiation – (Table 6.4). The control of airway homeostasis with the activation of the front line of defence mechanism in the early airway response to injury and cell cycle activities at later time points might have an association with the induction of rapid mechanism of repair of the airway in order to quickly regenerate a new epithelium layer. The down-play of such activities is in agreement with the morphological changes of the airway epithelium at this time point in which the airway sections showed evidence of mucosal haemorrhage at the damaged sites and flattened dedifferentiated cells within the transitional sites. Reduced cellular differentiation at this time point might in turn relate to the increase in proliferative activity observed from day 1 to day 7 post injury.

TIMEPOINTS	UP	DOWN
6 hours	GO:0006950~response to stress GO:0009611~response to wounding GO:0009605~response to external stimulus GO:0006954~inflammatory response GO:0009308~amine metabolic process	GO:0048518~positive regulation of biological process GO:0048522~positive regulation of cellular process GO:0032502~developmental process GO:0065007~biological regulation GO:0016043~cellular component organization and biogenesis
Day 1	GO:0007049~cell cycle GO:0022402~cell cycle process GO:0000278~mitotic cell cycle GO:0007067~mitosis GO:0000087~M phase of mitotic cell cycle	GO:0032502~developmental process GO:0048856~anatomical structure development GO:0030154~cell differentiation GO:0048869~cellular developmental process GO:0007275~multicellular organismal development
Day 3	GO:0000278~mitotic cell cycle GO:0007067~mitosis GO:0000087~M phase of mitotic cell cycle GO:0000279~M phase GO:0022403~cell cycle phase	GO:0065007~biological regulation GO:0009987~cellular process GO:0006950~response to stress GO:0050789~regulation of biological process GO:0032502~developmental process
Day 7	GO:0000279~M phase GO:0007067~mitosis GO:0000087~M phase of mitotic cell cycle GO:0000278~mitotic cell cycle GO:0006817~phosphate transport	GO:0008152~metabolic process GO:0044237~cellular metabolic process GO:0003012~muscle system process GO:0006936~muscle contraction GO:0044249~cellular biosynthetic process

Table 6.4: The list of GO term of biological process associate with the genes either up- or down-regulated following injury in each time point.
These genes are amongst the most significantly highly associated with the both set of genes in which at each time point, the most significant BP terms are listed at the top of the list.

In relation to the present study, twelve pathways associating with our list of genes were identified. These were cell cycle, ECM-receptor interaction, focal adhesion, small cell lung cancer, leukocyte transendothelial migration, cell communication, p53 signalling, TGF-beta signalling, hematopoietic cell lineage, toll-like receptor, complement and coagulation cascades and cytokine-cytokine receptor interaction pathways (Table 6.1). Of these only the TGF-beta signalling pathway was suggested to be involved early in the response to physical injury (6 hours) – (Table 6.3). Most of the pathways involved at this time point (6 hours) were those with a potential role in defence mechanisms and cellular migration. A wide spectrum of cellular functions such as proliferation, apoptosis, differentiation and migration are regulated by TGF-beta family members (Erjefalt JS et al., 1995). As all the genes involved in these activities or functions were down-regulated (as shown by GO term for BP) at this time point, it is conceivable that this represents a crucial transition point between the immediate needs of the site with respect to defence and damage limitation, and the activation of mechanisms geared towards subsequent repair. The involvement of hematopoietic cell lineage and leukocyte transendothelial migration at this time point suggest the early involvement of hematopoietic- and blood-derived cells in early cellular migration and induction of innate immunity. These cells undergoing differentiation process express a stage- and lineage-specific set of surface markers (Erjefalt JS et al., 1995). Therefore cellular stages are identified by the specific expression patterns of these genes. Small cell lung carcinoma (SCLC) is a highly aggressive neoplasm, and molecular mechanisms altered in SCLC include induced expression of oncogene, myelocytomatosis (MYC), and loss of tumorsuppressor genes, such as p53, phosphatase and tensin homolog (PTEN), retinoblastoma (RB), and fragile histidine triad (FHIT) (Erjefalt JS et al., 1995). The overexpression of MYC proteins in SCLC is largely a result of gene amplification (Erjefalt JS et al., 1995). Such over-expression leads to more rapid proliferation and loss of terminal differentiation (Erjefalt JS et al., 1995). The involvement of this pathway (SCLC) was only observed from 6 hours up to day 3 post injury.

Cell cycle and ECM-receptor interaction pathways were the only pathways that were involved in all time points following injury. The significant contribution of these

pathways especially for ECM-receptor interaction during the course of injury can be seen by the increasing significance associated with their involvement from 6 hours to day 7 post injury. ECM-receptor interaction pathways generally involve a complex mixture of structural and functional macromolecules and serves an important role in tissue and organ morphogenesis and in the maintenance of cell and tissue structure and function (Erjefalt JS et al., 1995). These interactions lead to a direct or indirect control of cellular activities such as adhesion, migration, differentiation, proliferation, and apoptosis – all clearly important in the repair of the injured airway wall.

The pathway analysis suggested that at days 3 and 7 post injury, most of the pathways associated with the genes regulated at these time points were involved in cells communication and re-structure of the damaged airway morphology. It is worth examining the potential interactions between pathways at this time. In this regard cell-matrix adhesions play essential roles in important biological processes including cell motility, cell proliferation, cell differentiation, regulation of gene expression and cell survival. At the cell-extracellular matrix contact points, specialized structures are formed and termed focal adhesions, where bundles of actin filaments are anchored to transmembrane receptors of the integrin family through a multi-molecular complex of junctional plaque proteins (Erjefalt JS et al., 1995). Whilst the cell-cell adherens junctions (AJs), the most common type of intercellular adhesions, are important for maintaining tissue architecture and cell polarity and can limit cell movement and proliferation (Erjefalt JS et al., 1995). At AJs, E-cadherin serves as an essential cell adhesion molecules (CAMs) (Erjefalt JS et al., 1995).

ECM-receptor and focal adhesion pathways shared similar gene groupings and thus were likely involved in similar mechanisms in response to airway injury and repair. Genes involved within these clusters included the families of laminin, collagen, integrin and thrombospondin genes. The involvement of the integrin gene family is reported in epithelial cell migration and adhesion to the basement membrane (Quaranta V, 1990) which also plays an important role in cell-cell and cell-extracellular matrix interactions. A review by (Bosman FT, 1993) clearly discussed

the roles of the integrin gene family in the interaction of the cell with the extracellular matrix in order to create and maintain tissue architecture. Integrins are expressed by epithelial cells on the basement membrane (Pilewski JM et al., 1997). In the context of the present study, following a bronchial brushing, there was evidence of loss of the basement membrane at the earlier time points at the site of the damaged epithelium (6 hours and day 1 post injury) and this loss presumably influenced the subsequent production of integrin and laminin. However, this study revealed that several of the integrin gene families presented on the array were significantly regulated at various time points after injury (Table 6.5). However only ITGA5 was significantly upregulated (fold change >2.0 on days 1, 3 and 7 post injury) whereas ITGA4 was marginally no changed (fold change ~0.8-0.85 at 6-hour and day 3 respectively) and ITGA8 was down-regulated (fold changes 0.27-0.49 through days 1-3 after injury). For the laminin gene, only laminin beta 1 (LAMB1) was found to be significantly upregulated from days 1-3 after injury (Chapter 6_Appendix 6.1_Significant at p less than 0.05_annotated on CD). This finding suggests that up-regulation of the laminin and both up-regulation of ITGA5 and down-regulation of ITGA8 genes are associated with regenerating a layer of epithelial cells at the damaged sites. These expression patterns may also relate to roles played by the regenerating epithelial cells themselves in regulating interactions with the extracellular matrix in order to maintain epithelial architecture (Bosman FT, 1993). The contrasting regulation of different integrins probably reflects the complexity of basement membrane structure and the fact that such structure may, in itself, show dynamic change during the course of repair. The various regulations of the integrin molecules in the earlier time points might associate with the production of this molecule by the migrating epithelial cells bordering the lesion or inflammatory response-related cells. The former assumption is based on the fact that the RNA isolated from the damaged epithelium tissue might also have contained elements from the transitional areas at both edges of the damaged site (Chapter 2; Figure 2.2 – the middle panel).

Probenames	TargetID	GeneSymbols	Timepoints	*p*-values	Fold Changes	Regulation (effect)	FDR
A_70_P026032	AF349458	ITGA4	0.25	0.010875	0.805598	No changed	0.036839
A_70_P026032	AF349458	ITGA4	3	0.048155	0.853837	No changed	0.11793
A_70_P029596	FE031898	ITGA5	1	0.013926	2.169972	Up-regulation	0.094726
A_70_P029596	FE031898	ITGA5	3	0.002357	2.812312	Up-regulation	0.026707
A_70_P029596	FE031898	ITGA5	7	0.0139	2.170562	Up-regulation	0.118498
A_70_P026036	DY499111	ITGA5	7	0.014525	1.209778	No changed	0.120687
A_70_P058166	FE032203	ITGA8	0.25	0.007551	0.374059	Down-regulation	0.030483
A_70_P058166	FE032203	ITGA8	1	0.005183	0.351454	Down-regulation	0.057167
A_70_P058166	FE032203	ITGA8	3	0.001063	0.268788	Down-regulation	0.017937
A_70_P058166	FE032203	ITGA8	7	0.038252	0.489729	Down-regulation	0.179046
A_70_P028307	DQ239659	LAMB1	1	3.33E-05	3.176645	Up-regulation	0.003823
A_70_P028306	DQ239659	LAMB1	1	1.04E-05	3.113726	Up-regulation	0.002025
A_70_P028307	DQ239659	LAMB1	3	0.000633	2.282111	Up-regulation	0.013828
A_70_P028306	DQ239659	LAMB1	3	0.00035	2.167268	Up-regulation	0.010254
A_70_P028307	DQ239659	LAMB1	7	0.018051	1.63732	No changed	0.132048
A_70_P028306	DQ239659	LAMB1	7	0.013052	1.579405	No changed	0.115428

Table 6.5: Expression of integrin and laminin genes during airway regeneration and repair following bronchial brushing.
Duplicates listed for specific time points reflect the fact that in some instances more than one gene-specific probe was present on the array and was therefore available to amplify the same gene. The listed time points represent those at which genes were found to be significantly expressed at $p<0.05$ as compared to the naive airway (day 0). The fold changes observed are relative to day 0 (naive). The significance changes observed from statistical analysis of microarray data are considered up- and/or down-regulated based on the values of fold changes between <0.5 (down-regulated), >2.0 (up-regulated) and 0.5>FC<2.0 (unchanged).

The consistent expression of the ITGA8 in the present study throughout the time course of injury might be associated with the involvement of basal cells in migration and cellular communication processes due to the fact that the expression of the beta ITGA8 was restricted to basal cells (Pilewski JM et al., 1997). A study by Pilewski JM et al., (1997) examined the effects of mechanical injury in an *in vivo* xenograft model of human bronchial epithelium. These authors reported that collagen-alpha 2 and 6 were produced by the supra-basal epithelial layer at the later phase of repair thus suggesting a role for these genes in airway migration (Pilewski JM et al., 1997). White SR et al., (1999) subsequently demonstrated that inhibition of these genes (collagen, integrin and laminin) led to impaired wound closure due to the lack of cellular interaction and communication (White SR et al., 1999). The up-regulation of the fibronectin gene (FN1) as early as 6 hours, and then at day 3 and day 7 post injury (Chapter 6_Appendix 6.1_Significant at p less than 0.05_annotated on CD) may suggest a role for this gene in this model system. A study by Han SW and Roman J, (2006) demonstrated that the fibronectin gene played a role in stimulating human airway epithelial cell growth and inhibiting apoptosis through activation of the NF-(kappa)B pathway (Han SW and Roman J, 2006).

The significant involvement of the cell cycle pathway from the day 1 up to day 7 post injury agreed with previous histopathological findings in which significant proliferative activity was appreciated at days 1, 3 and 7 in which at these time points (days 1 up to day 7), the most significant terms associated with the list of genes in these time points are involved in cell cycle regulation process, cell cycle phase and mitosis [refers to the (Table 6.4) for a summary of the up-regulated genes at the time points of days 1, 3 and 7]. As previously alluded the different density of the Ki67 staining seen on the histology sections supports the hypothesis that at any point in time cells were involved in different phases of the cell cycle. The involvement of complement and coagulation cascades pathway coincides with the evidence of mucosal haemorrhage (including fibrin, blood etc) at these time points (days 1, 3 and 7 post injury).

Clearly the process of regeneration and repair following injury involves different pathways with their respective involvement determined by the time point relative to injury. Respecting the diverse nature of the pathways concerned it is nonetheless clear that at any given time an equilibrium exists during this process that presumably reflects crosstalk and interaction both within, but presumably also between, pathways.

The existence of the specific pathways at each time point of injury and its progress towards repair is consistent with the histopathological finding. For instance, 6 hours after injury whilst there was evidence of mucosal loss and airway structure perturbation, neither an inflammatory response nor proliferative activity were prominent features. Most of changes observed at this time point were thought to be related to dedifferentiation and migration activities. In contrast, at the molecular level, more probes were significantly regulated ($p<0.05$) at this time point (4749 probes) as compared to any other time point. Most of the significantly regulated probes at this time point were in fact down-regulated. Whilst amongst those up-regulated were genes involved in inflammatory response, response to wounding and innate immunity and the majority of down-regulated genes were those involved in regulation of the biological process of the cells and function. The balance (between the proportions of up- and down-regulate genes) was re-dressed with time, presumably reflecting the activation of specific pathways and biological functions required in order to engineer the repair to the damaged epithelium.

The involvement of the TGF-beta signalling pathway early after injury warrants further discussion. TGF-beta1 induces epithelial-mesenchymal transition (EMT; the process by which differentiated epithelial cells undergo a phenotypic transition to mesenchymal cells (Kalluri R and Neilson EG, 2003)) in a Smad3-dependent manner in primary airway epithelial cells (AECs) (Hackett TL et al., 2009). EMT is characterized by loss of cell adhesion, repression of E-cadherin expression, and increased cell mobility and is a characteristic feature of cells undergoing proliferation. Members of the TGF-beta family of growth factors can initiate and maintain EMT in a variety of biological systems and different pathophysiological

contexts by activating major signalling pathways and transcriptional regulators integrated in extensive signalling networks (Zavadil J and Böttinger EP, 2005). It is tempting to speculate that the early molecular signature represents a "preparation phase" in which the airway recognises the type and extent of injury and prompts the activation of the pathways relevant to avoiding further damage to the airway brushing site. These include the migration of hematopoietic cell lineages and leukocyte transendothelial migration, cell cycle and toll-like receptor pathway activations.

A study of TGF-beta induced apoptosis in hepatocytes suggested that GADD45B functions as a critical upstream component in the apoptosis pathway induced by TGF-beta in the liver (Yoo J et al., 2003). This apoptosis process is required to eliminate the damaged or abnormal cells from normal tissue and thus the loss of the TGF-beta-induced apoptosis was associated with hepatocellular carcinoma (Sue SR et al., 1995). In the context of the present study, the TGF-beta (TGFB1) gene was significantly up-regulated at day 3 and day 7 post injury whilst GADD45 (GADD45A) gene was up-regulated at the earlier time points (6 hours and day 1) - (Chapter 6_Appendix 6.1_ Significant at p less than 0.05_annotated on CD). One of the functions of the TGFB1 is to stimulate ECM protein production (Lee JM et al., 2006). In the context of the present study, the down-regulation of the TGFB1 gene in the earlier time points 6 hours and day 1 post injury may reflect an early requirement to control or limit apoptosis. Indeed the expression of the genes within this TGF-beta pathway such as SMAD2, SMAD6 and SMAD7 genes were down-regulated at the 6-hour time point (Chapter 6_Appendix 6.1_ Significant at p less than 0.05_annotated on CD). The down-regulation of these genes might reduce or inhibit the activity of the apoptosis and cell cycle process (Appendix 6.12b – as attached). In contrast SMAD1 gene was found to be up-regulated at this time point and the activation of this gene might associate with the development of the airway structure following injury. A study by Chen C et al., (2005) suggested that airway epithelial cell proliferation and differentiation were inhibited when endogenous SMAD1 expression was knocked down. It was also been reported that SMAD1 expression was found in the smooth muscle and suggested that SMAD1-mediated signalling

may be required for smooth muscle cell growth or maintenance at a particular developmental stage (E16.5) (Chen C et al., 2005). Jeffery TK et al., (2005) reported that SMAD1-mediated bone morphogenetic protein 4 (BMP4) signalling is essential for myofibroblast differentiation into smooth muscle cells *in vitro* (Jeffery TK et al., 2005).

The down-regulation of TGF-beta at early time points contrasts with the observation that TGF-beta appears to stimulate the dedifferentiation process in which treating dedifferentiated rabbit renal proximal tubular cells (RPTC) with TGF-beta for 24-hour enhanced spindle morphology and vimentin expression compared with dedifferentiated cells treated with diluent (Hallman MA et al., 2008). In this study several markers were used to assess dedifferentiation activity in RPTC primary cultures following exposure to H_2O_2 (Hallman MA et al., 2008). Dedifferentiation was characterized using multiple markers, including vimentin expression, cell elongation, and the loss of N-cadherin and apical-basal polarity (Hallman MA et al., 2008). Epidermal growth factor receptor (EGFR) activation has been reported to be a critical step in dedifferentiation *in vitro* after oxidant injury and inhibition of EGFR expression stimulated a redifferentiation (Hallman MA et al., 2008). Zhuang S et al., (2005) similarly reported the use of molecular markers for dedifferentiation activity in RPTC primary cultures. In this latter study, it was revealed that p38 mediates epidermal growth factor (EGF) receptor activation after oxidant injury; that sarcoma (Src) activates Medicago MAP kinase gene 3 (MMK3), which, in turn, activates p38; and that the EGF receptor signalling pathway plays a critical role in RPTC dedifferentiation (Zhuang S et al., 2005).

The roles of the apoptosis in airway inflammation are clearly described by Monteseirín J, (2009). The apoptosis activity involves in eliminating of neutrophils during the allergic resolution and suppression of the apoptosis leads to prolong survival of these cells. Increases IL-8 secretion and inhibits neutrophil apoptosis and subsequent release of IL-8 may contribute to neutrophil inflammation in patients with atopic asthma (Monteseirín J, 2009). Apoptosis also may play an important role in COPD pathogenesis (Park JW et al., 2007). Increased numbers of apoptotic cells

can be detected in lung tissue and airways of human subjects with COPD and increased apoptosis in COPD can also imply defects in the normal physiological clearance of apoptotic cells (Park JW et al., 2007). However, the detection of apoptotic cells population is also used as a sign for cellular and tissue replacement following injury (Henson PM et al., 2006) in which the balance process between cell loss affected by the injury and cellular replication in order to restore its normal morphology.

The airway epithelium forms an important protective barrier to injuries induced by external stimuli, whether through smoke inhalation or aspiration of acid stomach contents, or through exposure to pollutants or infectious moeities in the atmosphere we breathe. It is important that the airway mounts a rapid response to such insult. Any breakdown in the epithelial barrier exposes the underlying architecture to further injury, leads to fluid loss through the epithelial surface and ultimately perturbs the innate defence of the lung.

Recently, it has been reported that the activation of TLRs by endogenous and exogenous ligands may play a central role in determining the balance between a state of "mucosal homeostasis," as is required for optimal organ function, and "mucosal injury," leading to mucosal inflammation and barrier breakdown (Sha Q et al., 2004, Gribar SC et al., 2008). It has also been suggested that the epithelial cells rapidly generate complement proteins for purposes of host defence when exposed to pathogens that express agonists for TLRs (Sha Q et al., 2004).

As highlighted by Gribar SC et al., 2008 not only is the epithelium able to respond to potentially dangerous microbial products, it also may sense endogenous molecules that are released during conditions of stress, hypoxia, or injury — so-called danger molecules that may play a critical role in the development of mucosal inflammation (Gribar SC et al., 2008). It assumes this important role at least in part through TLRs sensing the so-called "danger signals", also called damage-associated molecular patterns (DAMPs) and those ligands on the surface or interior of pathogen-associated molecular patterns (PAMPs). That these mechanisms are likely involved in the

normal process of epithelial repair following physical injury is evidenced by the early involvement of TLR pathways in the present study. Six TLR genes (TLR1, TLR5, TLR6, TLR8, TLR9 and TLR10) were significantly regulated (up, down or unchanged) - (Chapter 6_ Appendix 6.1_Significant at p less than 0.05_annotated on CD) and (Table 6.6). The array data suggested that TLR genes were significantly regulated at various stages of the injury and repair process. Only TLR 2 (at 6 hours) post injury was found to be up-regulated at any of the time point with majority of those genes were unchanged (Table 6.6). The majority of the TLRs demonstrated a time-dependent increase in expression from an initial low point at 6 hours (Chapter 6_ Appendix 6.1 on CD) especially TLR7 and TLR8 genes.

Gene Symbols	Fold changes			
	6 hours	Day 1	Day 3	Day 7
TLR1	0.53*	0.69	0.60*	0.85
TLR2	2.02*	1.65	1.56	1.40
TLR5	0.41*	0.55*	0.41*	0.68
TLR6	0.45*	0.67	0.61*	0.76
TLR7	0.71*	0.95	1.55*	1.87*
TLR8	0.71*	0.86	1.17	1.35*
TLR9	1.03	1.07	1.10*	1.04
TLR10	0.29*	0.44*	0.57	0.87

Table 6.6: Toll-like receptor genes with fold changes relative to naïve indicated over the time course of the study. * indicates the fold changes with significance *p*-values < 0.05
The observed changes are considered up- and/or down-regulated based on the values of fold changes between <0.5 (down-regulated), >2.0 (up-regulated) and 0.5>FC<2.0 (unchanged).

As shown in the appendix (Chapter 6_ Appendix 6.12d_Toll-like receptor signalling pathway – as attached) TLR2, TLR5 and TLR6 are involved in pathogen recognition and are capable of provoking rapid activation of innate immunity by inducing production of proinflammatory cytokines and up-regulation of costimulatory molecules via induction of myeloid differentiation factor-88 (MyD88), MAPK signalling and activation of NF-(kappa)B and AP-1. The activation of TLR2 is necessary to sufficiently stimulate NF-(kappa)B and MAPK pro-inflammatory signalling to recruit and activate polymorphonuclear leukocytes in the airway (Martin FJ and Prince AS, 2008) and TLR2 is also involved in regulating gap

junction intercellular communication in airway cells (Martin FJ and Prince AS, 2008). The quick activation of the NF-(kappa)B and MAPK signalling pathways within the TLR signalling pathway are important to provoke a rapid induction of innate immunity. The effect of such rapid induction is likely evidenced by the early up-regulation of IL6 and IL8 which in turn likely contributes to activating the cytokine-cytokine receptor interaction pathway. Up-regulation of IL6 is important to induce the acute-phase reaction which promotes the innate immune system and protects against tissue damage whilst the up-regulation of the IL8 gene is required for attracting the cells involved in defence mechanisms including neutrophils, macrophages and other leukocytes. An *in vitro* study involving an airway epithelial cell line and primary bronchial epithelial cell cultures demonstrated that IL8 expression (along with the expression of other TLR genes) significantly impacted on airway innate immune responses to pathogens (Sha Q et al., 2004).

There were interesting findings in terms of the regulation of the TLR genes observed in the array data as observed in the present study in which the contrasting roles of TLR2 compared to its component molecules (TLR1 and TLR6) during the early response to injury in which both TLR1 and TLR6 were significantly down-regulated whilst TLR2 been significantly up-regulated at 6 hours post injury as compared to naïve (time point 0). It is well-known that the activation of TLR2 requires homo- or heterodimerization with TLR1 or TLR6 (Chang S et al., 2007). By designing small interfering RNAs targeted to TLR2, TLR1, and TLR6, (Chang S et al., 2007) showed that knockdown of each of these receptors impaired pro- and anti-inflammatory cytokine activation by TLR-specific ligands in human embryonic kidney-TLR2 cells and in primary human macrophages (Chang S et al., 2007). The (Chapter 6_Appendix 6.12d – as attached) shows unique ability of TLR2 to form heteromers with TLR1 or TLR6 to mediate intracellular signalling in turn involved in activation of chemotactic and pro-inflammatory effects and apoptosis activity of the cells in response to injury. A study by (Farhat K et al., 2008) concluded that heterodimerization of TLR2 with TLR1 or TLR6 is important to enable the innate immune system to recognize the numerous, different structures of bacterial lipopeptides (LP) present in various pathogens (Chang S et al., 2007). In the context

of the present study, the mechanism of how these genes interacting to each other is unknown but one could be speculated is that the perturbation of these TLR genes might has induced the innate immune and inflammatory responses for downstream application in repair activity which clearly observed at the later time points (mucosal haemorrhage and proliferative activity).

In the present study, the early involvement of toll-like receptor pathways may play a role in regulating cells into the apoptosis stage. An *in vitro* model system of human first trimester trophoblasts revealed that the apoptosis induced by exposure to Gram-positive bacterial peptidoglycan (PDG) was found to be mediated by both TLR1 and TLR2 genes. TLR6 blocked the apoptosis pathway, activated trophoblast (NF-(kappa)B and triggered these cells to secrete the IL6 and IL8 necessary to regulate the inflammatory responses (Abrahams VM et al., 2008).

The p53 tumour suppressor protein is a multi-functional transcription factor that regulates cellular processes affecting proliferation, cell cycle checkpoints, and apoptosis and the consistent involvement of genes in this pathway indicates the importance of these processes in the airway epithelial response to injury. Apoptosis is well recognised in situations where cells fail to quickly repair following exposure to damage - whether at the level of the DNA or at the whole cell level. In the context of the present study, the brushing caused damage to cellular morphology and likely precipitated cell death. In addition, the activation of the cellular proliferation and cell cycle activities led to the generation of cells involved in the repair and regeneration process. Incumbent in such process would be the need to fine-tune and regulate the proportions of each cell type according to both the stage of injury and in relation to the ultimate repaired epithelial phenotype – hence the need to drive certain cells to apoptosis to accommodate this end (Tesfaigzi Y, 2003). Whilst apoptosis of neutrophils may benefit the host by helping the resolution of inflammation, apoptosis of epithelial cells may lead to further disruption of the epithelial barrier and lengthen the healing process. The activation of the p53 signalling pathway at day 1 up to day 3 post injury suggest the important roles of the apoptosis activity in regulating airway regeneration and repair. Clearly, an apoptotic balance may exist between the

different cell types involved and at the different stages of the repair process and it is probable that the derived information in this regard reflects the overall balance between the multitudes of cells involved. The genes involved in the p53 pathway clustered in two groups, one of which (CDKN1A, SFN and GADD45A) is involved in inhibiting cell cycle arrest whilst the second (PAI, TSP1 and GADD45A) is involved in the inhibition of angiogenesis and metastasis, DNA repair and damage prevention (Chapter 6_Appendix 6.9c_p53 signalling pathway – as attached). In this pathway, the involvement of the TGF-beta gene within this cluster is significant as this gene is known to be involved in controlling cellular proliferation and differentiation through inhibiting epithelial cell proliferation and induction of apoptosis (Yoo J et al., 2003).

The expression of the cell cycle-related genes early on in the response to injury was found to be associated with the regulation of the p53 signalling pathway that inhibits cellular senescence and cell cycle arrest. The activation of the cell cycle pathway was consistent with the evidence of cell proliferation which peaked at day 3 post injury. Most of the genes involved in the cell cycle pathway at day 1 post injury were the genes activated in all phases of the cell cycle stage of DNA replication (S-phase) and mitosis (M). In the overall analysis incorporating data from all time points, the GO term for mitosis was the most significant term associated with our list of genes. In fact, at every time point, this activity was among the most significant term associating with these significantly expressed genes. It is probable that the mitotic activity is essential in generating new daughter cells to regenerate the epithelium.

Most of the up-regulated cell cycle-related genes observed in this study were associated with the G_2 and M phases of the cell cycle – the phases during which protein synthesis and mitosis take place. This was consistent with a transcriptomic analysis of the response of the healthy human airway to bronchial brushing in which, at day 7 post injury, there was increased levels of many late stage cell cycle genes (G_2 and M phases). It was suggested that the involvement of these genes was consistent with the morphological assessment of the airway epithelium at day 7 post

injury when there was evidence of complete coverage of the brushed area (Heguy A et al., 2007).

6.6 CONCLUSION OF THE CHAPTER

By application of functional annotation and analysis tools contained within DAVID the biological functions of the groups of genes that were significantly regulated in our study were assessed. GO terms for the analogous human genes were established and used in the analysis to predict the biological process in the sheep model. The appearance of the significant GO terms identified in this study is dependent on the stages or epithelial repair involving inflammatory response, cellular migration, extracellular matrix activities, differentiation, proliferation, cellular development, cell cycle activities, cellular adhesion, apoptosis and mitosis. During the airway regeneration different KEGG pathways were demonstrated to be involved at specific times during the process of repair. Such data will provide a foundation upon which future studies can exploit the increasing availability of annotated genomic information pertaining to large animal model systems and will potentially prove useful in linking the processes seen during normal repair with those seen in conjunction with more complex models of airway epithelial damage such as those relating to asthma, COPD or cystic fibrosis phenotypes.

CHAPTER 7

GENERAL DISCUSSION

7.1 Summary of the results

The work presented in this thesis was aimed towards understanding the mechanisms involved in airway regeneration and repair in the sheep. This study demonstrates, for the first time, that bronchial brushing in anaesthetised animals provides a means of physical injury that is well tolerated and quickly repaired. Histopathology of brushed areas indicated that the injury completely denuded the brushed area of epithelium and indeed removed the underlying basement membrane and perturbed the deeper structures. Initial studies indicated that the major part of the repair process was complete within the first week after injury and further confirmed the utility of following changes over time involving multiple brush sites in the same animal.

The general process of repair appeared to broadly agree with that previously characterised in small animal tracheal injury systems - such repair being characterised by dedifferentiation of marginal epithelium, migration, proliferation and re-differentiation. However this model extends the value of such basic research in examining the repair process in the bronchi themselves, the site where this technique is commonly applied in clinical and veterinary clinical respiratory medicine, in a species with bronchial airway wall architecture comparable to that of human. With proliferation being a key part of repair these studies provide, for the first time, information about the dynamics, location and extent of proliferative activity in the more distal airway epithelium in large animals.

Certain limitations exist that confound the value of such large animal systems in agricultural species. Notably very few antibodies are available to label cells of interest in order to trace their function and roles during the repair process. As a consequence this study sought to define the molecular characteristics of the repair process. Towards this end semi-quantitative analysis of RT-PCR assays were

developed to probe expression profiles of genes that are believed to associate with specific cell types such as basal and ciliated cells. The genes that were considered in this way were beta-catenin, beta-tubulin, keratins 5 and 14, and FoxJ1. This analysis revealed that whilst the expression profiles of the studied genes did indeed vary over the course of the repair process the extent of within- and between-animal variability in this regard would likely erode the potential ability of this approach to glean clear and definitive statements regarding the likely pathways and mechanisms underlying the different stages of repair. Such ability would of course further be influenced by the nature and extent of gene selection – hinged as it was, and would be, on the available literature, but as a consequence biased, and less likely to pick up species-specific differences. Thoughts turned to the potential utility of microarray expression profiling using sheep microarrays made recently commercially available as a means of characterising the mechanisms underlying the repair process.

Microarray hybridisation generally requires a high quality RNA sample to meet quality standards for reproducibility and reliability of the microarray data and generally involves optimisation of sampling protocols. This reality was quickly appreciated in the context of samples derived from the airway wall and efforts were directed towards optimising the sampling protocols.

The experimental protocol that was designed to evaluate the time-dependent changes in gene expression concordant with histopathological features involved eight sheep. In order to facilitate evaluation of airways at four time points after injury (6 hours, days 1, 3 and 7) sheep were allocated to one of four possible sampling schedules (6 hours, days 1 and 3; 6 hours, days 1 and 7; 6 hours, days 3 and 7; days 1, 3 and 7) and data collected from the brushed areas at each time point was compared to spatially disparate unbrushed (naive) control airways within the same lung. The complexity of this experimental design reflected the need to optimise the data return to facilitate meaningful statistical appraisal whilst minimising ethical cost. This study identified more than 13,000 probes which were significantly expressed at various times during the repair process and amongst those were genes and pathways that appeared to correlate in a biologically plausible sense with the pathophysiology of

the wound repair process. One significant limitation to the use of such commercial microarrays was the relative lack of annotation associated with the arrayed ovine genes. With full-scale annotation beyond the scope of the present study a compromise was necessary in order to progress with the underlying objectives of this research – namely that a subset of genes would be annotated. To this end, the experimental analysis of microarray data identified for each gene and time point in expression with associated significance level. And decision was made to select only those annotated genes at a significance level less the 5% for further analysis. Out of 13, 495 probes significantly regulated at $p<0.05$, 6420 probes managed to be annotated and were selected for downstream application of biological and pathway analyses.

In order to better understand the biological function of the significantly expressed genes a further analysis based on their gene ontologies was completed. Gene ontology (GO) terms associate genes with relevant functions (biological process, cellular component and molecular function) and analysis seeks to use this information to piece together pathways and biology underlying complex expression patterns. The repair process observed at the histopathological level was consistent with the biological function proposed by the analysis tools and involved inflammatory and innate immunity responses, early phase (6 hours) migration, cellular proliferation, ECM function, and the activation of cell cycle processes, apoptosis and cellular re-differentiation at later time points. Several pathways were found to be associated with the phases of repair in our study. These were identified as cell cycle, ECM-receptor interaction, focal adhesion, small cell lung cancer, leukocyte transendothelial migration, cell communication, p53 signalling pathway, TGF-beta signalling pathway, hematopoietic cell lineage, toll-like receptor signalling pathway, complement and coagulation cascades and cytokine-cytokine receptor pathways. Such studies indicate the facility of this model system to dissect and understand the events that are associated with airway epithelial repair. Potential naturally exists to carry this forward and several avenues of further work and investigation are worth discussing.

7.2 Future works

Two major limitations were identified in the context of this large animal system. The first concerns the availability of immunological markers to follow the involvement of specific cell types throughout the repair process and the second concerns the lack of annotation available for the currently available sheep commercial microarray. In the first regard it is fortunate that initiatives are underway (The Immunological Toolbox at http://www.immunologicaltoolbox.com/ and (Entrican G et al., 2009)) to develop the tools necessary to complete this analysis and the resource of tissues and tissue sections archived during the course of the present study will facilitate the ongoing assessment of this work as more and more relevant antibody sets become available. Equally, the question of annotation will likely be addressed in the near future by the commercial interests concerned but in the meantime the process of searching probe targets against the human genomic databases in order to identify likely analogues will presumably lend itself to an automated approach that will facilitate iteration to continually update and/or allow re-specification of selection criteria. In the context of the methodologies employed in the current study such selection would build upon the list of human analogues with Human Gene Symbols that could be input into the DAVID functional analysis software and extend the power of this analysis tool.

The wider relevance of this study extends to abnormal airway epithelial repair in the context of airway pathology such as is observed in asthma, COPD and cystic fibrosis. The nature of the pathogenesis associated with these diseases share a common theme in that repetitive insult is believed to occur. How such repetitive insult impacts on the normal repair process is currently unknown but is a question that would lend itself well to the model system developed in this thesis. In particular, it would be possible to re-brush epithelium at different phases of the repair process to examine the impact of such an approach on the underlying gene expression and evoked pathways. Perturbing the normal process of repair and regeneration by repeating (perhaps on several occasions) the brushing process on the repairing epithelium would facilitate the study of the way in which such repeated insult influences the normal transcriptional pattern of repair. Indeed it may be possible to identify the pathways that continue to operate (in time and extent) on the basis of the original injury, those

that emerge as unique to the chronic injury and those that deviate from their anticipated pattern or time course of regulation initiated on the basis of the original injury. Such a model has not hitherto been characterized at the molecular level - although the idea of repeated injury has, as previously discussed (Gordon RE and Lane BP, 1977), been followed through to investigate cell cycle in small animal models. Such studies would have real relevance to understanding fundamental processes in complex diseases and potentially provide therapeutic angles that could be exploited in the future.

Another area of specific interest concerns the source of cells that are involved in the re-epithelialisation of the denuded area. In addition to the marginal epithelial de-differentiation, migration and re-differentiation, cells from the sub mucosal gland regions may also play a role. It is tempting to speculate that continual re-brushing of the denuded area will prevent these, and the wound marginal, cells from achieving the network of cell contacts that may be necessary to provoke the re-differentiation process.

Ordinarily, posing mechanistic questions about the roles of specific genes or pathways involves the strategic development of transgenic or knockout mouse models and observing the impact of such strategy on the system under study. Whether, in future, the potential exists to bring such mechanistic studies into the realms of large animal studies remains to be determined but it is conceivable that the very precision associated with achieving the brushing injury itself will also lend itself to the targeted application and or expression of small interfering molecules such as RNAi or specific peptides in the region of interest. Naturally such a strategy hinges on the detailed characterisation of the molecular response - perhaps even on an individual animal basis. Such studies would be prefaced and paralleled by analysis of *in vitro* systems. Indeed the basis for such studies is well established and could, for example, follow the techniques described by (Wesley UV et al., 2007) which was used the SiRNA technique to study the roles of dual oxidase 1 (Duox1) within the tracheobronchial epithelium in airway epithelial cell migration and repair following

injury *in vitro* culture using human primary tracheobronchial epithelium (HTBE) cell cultures.

Further refinement on our established technique could incorporate analysis of the brushed cells themselves. The molecular aspects of the associated methodology are well characterised within the sheep model and application of such would afford a means of following progress in airway epithelial repair without sacrifice of the animal. In any case the comparison of gene expression changes in the harvested cell population as compared with the airway wall itself would afford useful insight into the contribution of the epithelial surface to the observed gene expression changes in the wall itself.

Finally, and following on from the necessary groundwork involved in annotating the sheep array the potential exists to extend the analysis of our data through the application of new software tools as they become available. One such current example is a free programme called BioLayout Express^{3D} (Theocharidis A et al., 2009, Freeman TC et al., 2007). This programme capitalises on the remarkable visual processing capabilities of the human brain to analyze patterns and images and directs this capability towards assessing large scale bioinformatic data through representing the same in a highly visual context – in both 2- and 3-dimensional formats. Such approaches have proven useful to study gene expression data and networks within a wide variety of biological relationship systems (Freeman TC et al., 2007) and would lend themselves well to the further analysis of the datasets accumulated in the context of the present study. The networks consist of nodes representing transcripts or genes connected by virtue of their expression profile similarity across multiple conditions (this includes the individual sheep ID and the time points). The similarity between individual expression profiles may be determined by one of a number of possible statistical techniques, including Pearson and Spearman correlation coefficients. Thus networks can be constructed by connecting transcripts (nodes) by edges that infer varying degrees of co-expression based on an arbitrary correlation threshold thus establishing relationships between genes based on correlation of expression (Freeman TC et al., 2007, Theocharidis A et al., 2009). By being able to

assess the sheep-specific expression patterns of those genes at the various time points, both the fundamental aspects of the process conserved between sheep could potentially be identified, as well as the processes that contribute to the variation seen between sheep. We hypothesize that by using this type of approach it may be possible to broaden the scope of the microarray dataset by mapping the molecular networks and pathways that exist between genes and as such could establish the intra-relationship between genes within clusters and also the inter-relationship of the genes lying in separate clusters in order to improve our understanding of the behaviour of the complex networks that underlay the repair mechanism in the airway epithelium.

REFERENCES

(2008) World Health Statistics. *World Health Organisation*.

ABE S, BOYER C, LIU X, WEN FQ, KOBAYASHI T, FANG Q, WANG X, HASHIMOTO M, SHARP JG & RENNARD SI (2004) Cells derived from the circulation contribute to the repair of lung injury. *Am J Respir Crit Care Med. ,* 170**,** 1158-1163.

ABRAHAM WM (2008) Modeling of asthma, COPD and cystic fibrosis in sheep. *Pulm Pharmacol Ther,* 21**,** 743-54.

ABRAHAM WM, DELEHUNT JC, YERGER L & MARCHETTE B (1983) Characterization of a late phase pulmonary response after antigen challenge in allergic sheep. *Am Rev Respir Dis,* 128**,** 839-44.

ABRAHAM WM, WATSON H, SCHNEIDER A, KING M, YERGER L & SACKNER MA (1981) Noninvasive ventilatory monitoring by respiratory inductive plethysmography in conscious sheep. *J Appl Physiol,* 51**,** 1657-61.

ABRAHAMS VM, ALDO PB, MURPHY SP, VISINTIN I, KOGA K, WILSON G, ROMERO R, SHARMA S & MOR G (2008) TLR6 modulates first trimester trophoblast responses to peptidoglycan. *J Immunol. ,* 180**,** 6035-6043.

ADAM GIANGRECO, SUSAN D. REYNOLDS & STRIPP, B. R. (2002) Terminal Bronchioles Harbor a Unique Airway Stem Cell Population That Localizes to the Bronchoalveolar Duct Junction. *American Journal of Pathology,* 161**,** 173–182.

AGILENT TECHNOLOGIES INC GeneSpring GX Version 10.

AHMED T, EYRE P, JANUSZKIEWICZ AJ & WANNER A (1980) Role of H1- and H2-receptors in airway reactions to histamine in conscious sheep. *J Appl Physiol,* 49**,** 826-33.

AIGNER T & KIM HA (2002) Apoptosis and cellular vitality: issues in osteoarthritic cartilage degeneration. *Arthritis Rheum. ,* 46**,** 1986-96.

AL-SHAHROUR F, DÍAZ-URIARTE R & DOPAZO J (2004) FatiGO: a web tool for finding significant associations of Gene Ontology terms with groups of genes. *Bioinformatics. ,* 20**,** 578-580.

ALLISON, D. B., GADBURY. G.L., HEO, M., FERNANDEZ, J. R., LEE, C.-K., PROLLA, T. A. & WEINDRUCH R (2002) *A mixture model approach for the analysis of microarray gene expression data*.

ALPARD SK, ZWISCHENBERGER JB, TAO W, DEYO DJ, TRABER DL & BIDANI A (2000) New clinically relevant sheep model of severe respiratory failure secondary to combined smoke inhalation/cutaneous flame burn injury. *Crit Care Med,* 28**,** 1469-76.

ANDERSON DJ, GAGE FH & WEISSMAN IL (2001) Can stem cells cross lineage boundaries? *Nat Med,* 7**,** 393-395.

ANDREWS GK (2000) Regulation of metallothionein gene expression by oxidative stress and metal ions. *Biochem Pharmacol. ,* 59**,** 95-104.

ANTHONY B, DAVID HE, JAMES CH, ALAN LJ, RODNEY KL, MARA SL, JAMES M, DONALD MM, WAYNE AM, MITSUSHI O, RODGER JP, PETER DP, ROBERT RS, HARM MT, ELIZABETH MW & DEBORAH Y (1994) Proposed nomenclature for quantifying subdivisions of the bronchial wall. *J Appl Physiol,* 77**,** 1011-1014.

AOKI T, MATSUMOTO Y, HIRATA K, OCHIAI K, OKADA M, ICHIKAWA K, SHIBASAKI M, ARINAMI T, SUMAZAKI R & NOGUCHI E (2009) Expression profiling of genes related to asthma exacerbations. *Clin Exp Allergy,* 39**,** 213-21.

ASAKURA A & RUDNICKI MA (2002) Side population cells from diverse adult tissues are capable of in vitro hematopoietic differentiation. *Experimental Hematology,* 30**,** 1339–1345.

ASHBURNER M, BALL CA, BLAKE JA, BOTSTEIN D, BUTLER H, CHERRY JM, DAVIS AP, DOLINSKI K, DWIGHT SS, EPPIG JT, HARRIS MA, HILL DP, ISSEL-TARVER L, KASARSKIS A, LEWIS S, MATESE JC, RICHARDSON JE, RINGWALD M, RUBIN GM & SHERLOCK G (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet. ,* 25**,** 25-29.

BARROW RE, WANG CZ, COX RA & EVANS MJ (1992) Cellular sequence of tracheal repair in sheep after smoke inhalation injury. *Lung,* 170**,** 331-8.

BARROW RE, WANG CZ, EVANS MJ & HERNDON DN (1993) Growth factors accelerate epithelial repair in sheep trachea. *Lung,* 171**,** 335-44.

BATES JH, RINCON M & IRVIN CG (2009) Animal models of asthma. *Am J Physiol Lung Cell Mol Physiol,* 297**,** L401-10.

BECCI PJ, MCDOWELL EM & TRUMP BF (1978) The respiratory epithelium. II. Hamster trachea, bronchus, and bronchioles. *J Natl Cancer Inst. ,* 61**,** 551-61.

BEER HD, BITTNER M, NIKLAUS G, MUNDING C, MAX N, GOPPELT A & WERNER S (2005) The fibroblast growth factor binding protein is a novel interaction partner of FGF-7, FGF-10 and FGF-22 and regulates FGF activity: implications for epithelial repair. *Oncogene,* 24**,** 5269-5277.

BÉGIN R, ROLA-PLESZCZYNSKI M, MASSÉ S, LEMAIRE I, SIROIS P, BOCTOR M, NADEAU D, DRAPEAU G & BUREAU MA (1983a) Asbestos-induced lung injury in the sheep model: the initial alveolitis. *Environ Res,* 30**,** 195-210.

BÉGIN R, ROLA-PLESZCZYNSKI M, MASSÉ S, NADEAU D & DRAPEAU G (1983b) Assessment of progression of asbestosis in the sheep model by bronchoalveolar lavage and pulmonary function tests. *Thorax,* 38**,** 449-57.

BÉGIN R, ROLA-PLESZCZYNSKI M, SIROIS P, MASSE S, NADEAU D & BUREAU MA (1981) Sequential analysis of the bronchoalveolar milieu in conscious sheep. *J Appl Physiol,* 50**,** 665-71.

BEISSBARTH T & SPEED TP (2004) GOstat: find statistically overrepresented Gene Ontologies within a group of genes. *Bioinformatics. ,* 20**,** 1464-1465.

BELLUSCI S (2008) Lung stem cells in the balance. *Nat Genet.* , 40**,** 822-824.

BENJAMINI, Y. & HOCHBERG, Y. (1995) Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society,* 57**,** 289–300.

BERNS A (2005) Stem Cells for Lung Cancer? *Cell,* 121**,** 811-813.

BHATTACHARYA S & MARIANI TJ (2009) Array of hope: expression profiling identifies disease biomarkers and mechanism. *Biochem Soc Trans.* , 37**,** 855-862.

BISCHOF RJ, SNIBSON K, SHAW R & MEEUSEN EN (2003) Induction of allergic inflammation in the lungs of sensitized sheep after local challenge with house dust mite. *Clin Exp Allergy.* , 33**,** 367-75.

BISHOP AE (2004) Pulmonary epithelial stem cells. *Cell Prolif,* 37**,** 89-96.

BLATT EN, YAN XH, WUERFFEL MK, HAMILOS DL & BRODY SL (1999) Forkhead transcription factor HFH-4 expression is temporally related to ciliogenesis. *Am J Respir Cell Mol Biol.* , 21**,** 168-76.

BOERS JE, AMBERGEN AW & THUNNISSEN FB (1998) Number and proliferation of basal and parabasal cells in normal human airway epithelium. *Am J Respir Crit Care Med,* 157**,** 2000-2006.

BOERS JE, DEN BROK JL, KOUDSTAAL J, ARENDS JW & THUNNISSEN FB (1996) Number and proliferation of neuroendocrine cells in normal human airway epithelium. *Am J Respir Crit Care Med,* 154**,** 758-63.

BORRIE J & MITCHELL RM (1960) The sheep as an experimental animal in surgical science. *Br J Surg.* , 47**,** 435-45.

BORTHWICK DW, SHAHBAZIAN M, KRANTZ QT, DORIN JR & RANDELL SH (2001) Evidence for stem-cell niches in the tracheal epithelium. *Am. J. Respir. Cell Mol. Biol.,* 24**,** 662–670.

BOSMAN FT (1993) Integrins: cell adhesives and modulators of cell function. *Histochem J.* , 25**,** 469-77.

BOUSQUET J & KHALTAEV N (2007) Global surveillance, prevention and control of chronic respiratory diseases : a comprehensive approach. *World Health Organisation*.

BRASS DM, YANG IV, KENNEDY MP, WHITEHEAD GS, RUTLEDGE H, BURCH LH & SCHWARTZ DA (2008) Fibroproliferation in LPS-induced airway remodeling and bleomycin-induced fibrosis share common patterns of gene expression. *Immunogenetics,* 60**,** 353-69.

BROACKES-CARTER FC, MOUCHEL N, GILL D, HYDE S, BASSETT J & HARRIS A (2002) Temporal regulation of CFTR expression during ovine lung development: implications for CF gene therapy. *Hum Mol Genet,* 11**,** 125-31.

BTS (2001) Burden of Lung Disease. *British Thoracic Society,* 1st Edition.

BTS (2006) Burden of Lung Disease. *British Thoracic Society,* 2nd Edition.

CANNING BJ & WRIGHT JL (2008) Animal models of asthma and chronic obstructive pulmonary disease. *Pulm Pharmacol Ther,* 21**,** 695.

CARDOSO WV & LÜ J (2006) Regulation of early lung morphogenesis: questions, facts and controversies. *Development. ,* 133**,** 1611-24.

CARLA F. BENDER KIM, ERICA L. JACKSON, AMBER E. WOOLFENDEN, SHARON LAWRENCE, IMRAN BABAR, SINAE VOGEL, DENISE CROWLEY, RODERICK T. BRONSON & JACKS, T. (2005) Identification of Bronchioalveolar Stem Cells in Normal Lung and Lung Cancer. *Cell,* 121**,** 823–835.

CAVARRA E, FARDIN P, FINESCHI S, RICCIARDI A, DE CUNTO G, SALLUSTIO F, ZORZETTO M, LUISETTI M, PFEFFER U, LUNGARELLA G & VARESIO L (2009) Early response of gene clusters is associated with mouse lung resistance or sensitivity to cigarette smoke. *Am J Physiol Lung Cell Mol Physiol,* 296**,** L418-29.

CHANG S, DOLGANIUC A & SZABO G (2007) Toll-like receptors 1 and 6 are involved in TLR2-mediated macrophage activation by hepatitis C virus core and NS3 proteins. *J Leukoc Biol,* 82**,** 479-87.

CHAPMAN RW (2008) Canine models of asthma and COPD. *Pulm Pharmacol Ther,* 21**,** 731-42.

CHEN C, CHEN H, SUN J, BRINGAS P JR, CHEN Y, WARBURTON D & SHI W (2005) Smad1 expression and function during mouse embryonic lung branching morphogenesis. *Am J Physiol Lung Cell Mol Physiol,* 288**,** L1033-1039.

CHEN ZH, KIM HP, RYTER SW & CHOI AM (2008) Identifying targets for COPD treatment through gene expression analyses. *Int J Chron Obstruct Pulmon Dis,* 3**,** 359-70.

CHOMCZYNSKI P (1992) Solubilization in formamide protects RNA from degradation. *Nucleic Acids Res. ,* 20**,** 3791-2.

CHOW CW, HERRERA ABREU MT, SUZUKI T & DOWNEY GP (2003) Oxidative stress and acute lung injury. *Am J Respir Cell Mol Biol,* 29**,** 427-431.

CHUANG CY, CHANG CH & HUANG YL (2009) Thioredoxin mediates remodeling factors of human bronchial epithelial cells upon interaction with house dust mite-stimulated eosinophils. *Inhal Toxicol. ,* 21**,** 153-67.

COHEN L, E X, TARSI J, RAMKUMAR T, HORIUCHI TK, COCHRAN R, DEMARTINO S, SCHECHTMAN KB, HUSSAIN I, HOLTZMAN MJ, CASTRO M & THE NHLBI SEVERE ASTHMA RESEARCH PROGRAM (SARP) (2007) Epithelial cell proliferation contributes to airway remodeling in severe asthma. *Am J Respir Crit Care Med,* 176**,** 138-45.

COLEBATCH HJ & HALMAGYI DF (1961) Lung mechanics and resuscitation after fluid aspiration. *J Appl Physiol.,* 16**,** 684-96.

COLEBATCH HJ & HALMAGYI DF (1963) Effect of vagotomy and vagal stimulation on lung mechanics and circulation. *J Appl Physiol. ,* 18**,** 881-7.

COLLIE DD, MACALDOWIE CN, PEMBERTON AD, WOODALL CJ, MCLEAN N, HODGSON C, KENNEDY MW & MILLER HR (2001) Local lung responses following local lung challenge with recombinant lungworm antigen in systemically sensitized sheep. *Clin Exp Allergy. ,* 31**,** 1636-47.

COLLIE DD, MCLEAN N, SALLENAVE JM, BAKER A, BLUNDELL R, MILNE E, RHIND S & WOODALL C (2006) Local lung responses following endobronchial elastase and lipopolysaccharide instillation in sheep. *Int J Chron Obstruct Pulmon Dis. ,* 1**,** 189-99.

CONESA A, GÖTZ S, GARCÍA-GÓMEZ JM, TEROL J, TALÓN M & ROBLES M (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics. ,* 21**,** 3674-3676.

COPLAND I & POST M (2004) Lung development and fetal lung growth. *Paediatr Respir Rev.,* 5 S259-64.

COPOIS V, BIBEAU F, BASCOUL-MOLLEVI C, SALVETAT N, CHALBOS P, BAREIL C, CANDEIL L, FRASLON C, CONSEILLER E, GRANCI V, MAZIÈRE P, KRAMAR A, YCHOU M, PAU B, MARTINEAU P, MOLINA F & DEL RIO M (2007) Impact of RNA degradation on gene expression profiles: assessment of different methods to reliably determine RNA quality. *J Biotechnol. ,* 127**,** 549-59.

COSTA RH, KALINICHENKO VV & LIM L (2001) Transcription factors in mouse lung development and function. *Am J Physiol Lung Cell Mol Physiol,* 280**,** 823-838.

COULSON NM, JANUSZKIEWICZ AJ, DODD KT & RIPPLE GR (1989) The cardiorespiratory effects of diazepam-ketamine and xylazine-ketamine anesthetic combinations in sheep. *Lab Anim Sci,* 39**,** 591-7.

COULSON NM, JANUSZKIEWICZ AJ & RIPPLE GR (1991) Physiological responses of sheep to two hours anaesthesia with diazepam-ketamine. *Vet Rec,* 129**,** 329-32.

COX RA, BURKE AS, OLIVERAS G, ENKHBAATAR P, TRABER LD, ZWISCHENBERGER JB, JESCHKE MG, SCHMALSTIEG FC, HERNDON DN, TRABER DL & HAWKINS HK (2005) Acute bronchial obstruction in sheep: histopathology and gland cytokine expression. *Exp Lung Res. ,* 31**,** 819-37.

COX RA, JACOB S, OLIVERAS G, MURAKAMI K, ENKHBAATAR P, TRABER L, SCHMALSTIEG FC, HERNDON DN, TRABER DL & HAWKINS HK (2009) Pulmonary expression of nitric oxide synthase isoforms in sheep with smoke inhalation and burn injury. *Exp Lung Res,* 35**,** 104-118.

CRAIG NM, MILLER HR, SMITH WD & KNIGHT PA (2007) Cytokine expression in naïve and previously infected lambs after challenge with Teladorsagia circumcincta. *Vet Immunol Immunopathol.,* 120**,** 47-54.

CRAWLEY, M. J. (2007) *The R book,* Imperial College London at Silwood Park, UK, John Wiley and Sons, Ltd.

DABBAGH K, TAKEYAMA K, LEE HM, UEKI IF, LAUSIER JA & NADEL JA (1999) IL-4 induces mucin gene expression and goblet cell metaplasia in vitro and in vivo. *J Immunol. ,* 162**,** 6233-6237.

DAKIR EL, FEIGENBAUM L & LINNOILA RI (2008) Constitutive expression of human keratin 14 gene in mouse lung induces premalignant lesions and squamous differentiation. *Carcinogenesis,* 29**,** 2377-84.

DAVIES LA, MCLACHLAN G, SUMNER-JONES SG, FERGUSON D, BAKER A, TENNANT P, GORDON C, VRETTOU C, BAKER E, ZHU J, ALTON EW, COLLIE DD, PORTEOUS DJ, HYDE SC & GILL DR (2008) Enhanced lung gene expression after aerosol delivery of concentrated pDNA/PEI complexes. *Mol Ther. ,* 16**,** 1283-90.

DE LANGHE SP & REYNOLDS SD (2008) Wnt signaling in lung organogenesis. *Organogenesis. ,* 4**,** 100-8.

DEMAYO F, MINOO P, PLOPPER CG, SCHUGER L, SHANNON J & TORDAY JS (2002) Mesenchymal-epithelial interactions in lung development and repair: are modeling and remodeling the same process? *Am J Physiol Lung Cell Mol Physiol,* 283**,** L510-L517.

DEMOLY P, SIMONY-LAFONTAINE J, CHANEZ P, PUJOL JL, LEQUEUX N, MICHEL FB & BOUSQUET J (1994) Cell proliferation in the bronchial mucosa of asthmatics and chronic bronchitics. *Am J Respir Crit Care Med,* 150**,** 214-7.

DENNIS G JR, SHERMAN BT, HOSACK DA, YANG J, GAO W, LANE HC & LEMPICKI RA (2003) DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol. ,* 4**,** P3.

DONIGER SW, SALOMONIS N, DAHLQUIST KD, VRANIZAN K, LAWLOR SC & CONKLIN BR (2003) MAPPFinder: using Gene Ontology and GenMAPP to create a global gene-expression profile from microarray data. *Genome Biol. ,* 4**,** R7.

DUNCAN W. BORTHWICK, MARIAM SHAHBAZIAN, Q. TODD KRANTZ, JULIA R. DORIN, A. & RANDELL, S. H. (2001) Evidence for Stem-Cell Niches in the Tracheal Epithelium. *Am. J. Respir. Cell Mol. Biol.,* 24**,** 662–670.

EMERSON M, RENWICK L, TATE S, RHIND S, MILNE E, PAINTER HA, BOYD AC, MCLACHLAN G, GRIESENBACH U, CHENG SH, GILL DR, HYDE SC, BAKER A, ALTON EW, PORTEOUS DJ & COLLIE DD (2003) Transfection efficiency and toxicity following delivery of naked plasmid DNA and cationic lipid-DNA complexes to ovine lung segments. *Mol Ther. ,* 8**,** 646-53.

EMURA, M. (2002) Stem cells of the respiratory tract. *Paediatr. Respir. Rev,* 3**,** 36-40.

EMURA M (2002) Stem cells of the respiratory tract. *Paediatr. Respir. Rev,* 3**,** 36-40.

ENGELHARDT JF (2001) Stem cell niches in the mouse airway. *Am. J. Respir. Cell Mol. Biol.,* 24**,** 649-652.

ENGELHARDT JF, SCHLOSSBERG H, YANKASKAS JR & DUDUS L (1995) Progenitor cells of the adult human airway involved in submucosal gland development. *Development* 121**,** 2031-2046.

ENTRICAN G, LUNNEY JK, RUTTEN VP & BALDWIN CL (2009) A current perspective on availability of tools, resources and networks for veterinary immunology. *Vet Immunol Immunopathol. ,* 128**,** 24-9.

ERJEFALT JS, ERJEFALT I, SUNDLER F & PERSSON CG (1995) Effects of topical budesonide on epithelial restitution in vivo in guinea pig trachea. *Thorax* 50**,** 785-792.

ERJEFÄLT JS, ERJEFÄLT I, SUNDLER F & PERSSON CG (1995) In vivo restitution of airway epithelium. *Cell Tissue Res. ,* 281**,** 305-16.

ERJEFÄLT JS & PERSSON CG (1997) Airway epithelial repair: breathtakingly quick and multipotentially pathogenic. *Thorax. ,* 52**,** 1010-2.

ERJEFÄLT JS, SUNDLER F & PERSSON CG (1997) Epithelial barrier formation by airway basal cells. *Thorax.,* 52**,** 213-7.

EVANS MJ, COX RA, SHAMI SG & PLOPPER CG (1990) Junctional adhesion mechanisms in airway basal cells. *Am J Respir Cell Mol Biol. ,* 3**,** 341-7.

EVANS MJ, COX RA, SHAMI SG, WILSON B & PLOPPER CG (1989) The role of basal cells in attachment of columnar cells to the basal lamina of the trachea. *Am J Respir Cell Mol Biol. ,* 1**,** 463-9.

EVANS MJ & PLOPPER CG (1988) The role of basal cells in adhesion of columnar epithelium to airway basement membrane. *Am Rev Respir Dis. ,* 138**,** 481-3.

FARHAT K, RIEKENBERG S, HEINE H, DEBARRY J, LANG R, MAGES J, BUWITT-BECKMANN U, RÖSCHMANN K, JUNG G, WIESMÜLLER KH & ULMER AJ (2008) Heterodimerization of TLR2 with TLR1 or TLR6 expands the ligand spectrum but does not lead to differential signaling. *J Leukoc Biol,* 83**,** 692-701.

FARRAGHER SM, TANNEY A, KENNEDY RD & PAUL HARKIN D (2008) RNA expression analysis from formalin fixed paraffin embedded tissues. *Histochem Cell Biol. ,* 130**,** 435-45.

FLEIGE S & PFAFFL MW (2006) RNA integrity and the effect on the real-time qRT-PCR performance. *Mol Aspects Med,* 27**,** 126-39.

FLEIGE S, WALF V, HUCH S, PRGOMET C, SEHM J & PFAFFL MW (2006) Comparison of relative mRNA quantification models and the impact of RNA integrity in quantitative real-time RT-PCR. *Biotechnol Lett. ,* 28**,** 1601-13.

FREEMAN TC, GOLDOVSKY L, BROSCH M, VAN DONGEN S, MAZIÈRE P, GROCOCK RJ, FREILICH S, THORNTON J & ENRIGHT AJ (2007)

Construction, visualisation, and clustering of transcription networks from microarray expression data. *PLoS Comput Biol.* , 3**,** 2032-2042.

GERDES J, LEMKE H, BAISCH H, WACKER HH, SCHWAB U & STEIN H (1984) Cell cycle analysis of a cell proliferation-associated human nuclear antigen defined by the monoclonal antibody Ki-67. *J Immunol.* , 133**,** 1710-1715.

GERSON C, SABATER J, SCURI M, TORBATI A, COFFEY R, ABRAHAM JW, LAUREDO I, FORTEZA R, WANNER A, SALATHE M, ABRAHAM WM & CONNER GE (2000) The lactoperoxidase system functions in bacterial clearance of airways. *Am J Respir Cell Mol Biol,* 22 665-71.

GIANGRECO A, ARWERT EN, ROSEWELL IR, SNYDER J, WATT FM & STRIPP BR (2009) Stem cells are dispensable for lung homeostasis but restore airways after injury. *Proc Natl Acad Sci U S A.,* 106**,** 9286-9291.

GIANGRECO A, REYNOLDS SD & STRIPP BR (2002) Terminal bronchioles harbor a unique airway stem cell population that localizes to the bronchoalveolar duct junction. *Am J Pathol,* 161**,** 173–182.

GORDON RE & LANE BP (1976) Regeneration of rat tracheal epithelium after mechanical injury. II. Restoration of surface integrity during the early hours after injury. *Am Rev Respir Dis.,* 113**,** 799-807.

GORDON RE & LANE BP (1977) Cytokinetics of rat tracheal epithelium stimulated by mechanical trauma. *Cell Tissue Kinet,* 10**,** 171-181.

GRIBAR SC, RICHARDSON WM, SODHI CP & HACKAM DJ (2008) No longer an innocent bystander: epithelial toll-like receptor signaling in the development of mucosal inflammation. *Mol Med.* , 14**,** 645-59.

HACKETT BP, BRODY SL, LIANG M, ZEITZ ID, BRUNS LA & GITLIN JD (1995) Primary structure of hepatocyte nuclear factor/forkhead homologue 4 and characterization of gene expression in the developing respiratory and reproductive epithelium. *Proc Natl Acad Sci U S A.* , 92**,** 4249-53.

HACKETT TL, WARNER SM, STEFANOWICZ D, SHAHEEN F, PECHKOVSKY DV, MURRAY LA, ARGENTIERI R, KICIC A, STICK SM, BAI TR & KNIGHT DA (2009) Induction of epithelial-mesenchymal transition in primary airway epithelial cells from patients with asthma by transforming growth factor-beta1. *Am J Respir Crit Care Med,* 180**,** 122-33.

HALLMAN MA, ZHUANG S & SCHNELLMANN RG (2008) Regulation of dedifferentiation and redifferentiation in renal proximal tubular cells by the epidermal growth factor receptor. *J Pharmacol Exp Ther,* 325**,** 520-8.

HALMAGYI DF & COLEBATCH HJ (1961a) Cardiorespiratory effects of experimental lung embolism. *J Clin Invest.* , 40**,** 1785-96.

HALMAGYI DF & COLEBATCH HJ (1961b) Some cardiorespiratory parameters in anesthetized sheep. *J Appl Physiol,* 16.

HAN SW & ROMAN J (2006) Fibronectin induces cell proliferation and inhibits apoptosis in human bronchial epithelial cells: pro-oncogenic effects mediated by PI3-kinase and NF-kappa B. *Oncogene.* , 25**,** 4341-9.

HANSEL NN & DIETTE GB (2007) Gene expression profiling in human asthma. *Proc Am Thorac Soc.* , 4**,** 32-6.

HARDILLO J, VANCLOOSTER C & DELAERE PR (2001) An investigation of airway wound healing using a novel in vivo model. *Laryngoscope* 111**,** 1174-1182.

HARKEMA JR, MARIASSY A, GEORGE JS, HYDE DM & PLOPPER CG (1991) *Epithelial cells in the conducting airways: A species comparison,* New York, Basel & Hongkong, Marcel Dekker, Inc.

HARVEY BG, HEGUY A, LEOPOLD PL, CAROLAN BJ, FERRIS B & CRYSTAL RG (2007) Modification of gene expression of the small airway epithelium in response to cigarette smoking. *J Mol Med.* , 85**,** 39-53.

HEGUY A, HARVEY BG, LEOPOLD PL, DOLGALEV I, RAMAN T & CRYSTAL RG (2007) Responses of the human airway epithelium transcriptome to in vivo injury. *Physiol Genomics.* , 29**,** 139-48.

HEGUY A, HARVEY BG, O'CONNOR TP, HACKETT NR & CRYSTAL RG (2003) Sampling-dependent up-regulation of gene expression in sequential samples of human airway epithelial cells. *Mol Med.* , 9**,** 200-8.

HENSON PM, VANDIVIER RW & DOUGLAS IS (2006) Cell death, remodeling, and repair in chronic obstructive pulmonary disease? *Proc Am Thorac Soc.* , 3**,** 713-717.

HILDING AC (1965) Regeneration of respiratory epithelium after minimal surface trauma. *Ann Otol Rhinol Laryngol,* 74**,** 903-14.

HODGE S, HODGE G, FLOWER R, REYNOLDS PN, SCICCHITANO R & HOLMES M (2002) Up-regulation of production of TGF-beta and IL-4 and down-regulation of IL-6 by apoptotic human bronchial epithelial cells. *Immunol Cell Biol.* , 80**,** 537-43.

HOFFMAN A, HENDERSON AC, TSAI L & INGENITO E (2003) Physiologic responses of sheep to two different methods of papain exposure. *Inhal Toxicol* 15**,** 761-780.

HOLLE, S. & BIRTLES, M. (1990) An immunocytochemical method for studying patterns of cell proliferation in the wool follicle. *New Zealand Vet J,* 38**,** 89-93.

HONG KU, REYNOLDS SD, GIANGRECO A, HURLEY CM & STRIPP BR (2001) Clara cell secretory protein-expressing cells of the airway neuroepithelial body microenvironment include a label-retaining subset and are critical for epithelial renewal after progenitor cell depletion. *Am J Respir Cell Mol Biol,* 24**,** 671-681.

HONG KU, REYNOLDS SD, WATKINS S, FUCHS E & STRIPP BR (2004a) Basal cells are a multipotent progenitor capable of renewing the bronchial epithelium *Am J Pathol,* 164**,** 577-588.

HONG KU, REYNOLDS SD, WATKINS S, FUCHS E & STRIPP BR (2004b) In vivo differentiation potential of tracheal basal cells: evidence for multipotent and unipotent subpopulations. *Am J Physiol Lung Cell Mol Physiol,* 286**,** 643–649.

HUANG DA W, SHERMAN BT & LEMPICKI RA (2008) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res. ,* 37.

HUANG DA W, SHERMAN BT & LEMPICKI RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc.,* 4**,** 44-57.

HUANG DA W, SHERMAN BT, TAN Q, COLLINS JR, ALVORD WG, ROAYAEI J, STEPHENS R, BASELER MW, LANE HC & LEMPICKI RA (2007) The DAVID Gene Functional Classification Tool: a novel biological module-centric algorithm to functionally analyze large gene lists. *Genome Biol.,* 8**,** R183.

HUR T, SQUIBB K, COSMA G, HOROWITZ S, PIEDBOEUF B, BOWSER D & GORDON T (1999) Induction of metallothionein and heme oxygenase in rats after inhalation of endotoxin. *J Toxicol Environ Health A. ,* 56**,** 183-203.

HUTCHISON AA, BERNARD GR, SNAPPER JR & BRIGHAM KL (1984) Effect of aerosol histamine on lung lymph in awake sheep. *J Appl Physiol,* 56**,** 1090-8.

HUTCHISON AA, BRIGHAM KL & SNAPPER JR (1982) Effect of histamine on lung mechanics in sheep. Comparison of aerosol and parenteral administration. *Am Rev Respir Dis,* 126**,** 1025-9.

IMBEAUD S, GRAUDENS E, BOULANGER V, BARLET X, ZABORSKI P, EVENO E, MUELLER O, SCHROEDER A & AUFFRAY C (2005) Towards standardization of RNA quality assessment using user-independent classifiers of microcapillary electrophoresis traces. *Nucleic Acids Res. ,* 33**,** e56.

IMGRUND M, GRÖNE E, GRÖNE HJ, KRETZLER M, HOLZMAN L, SCHLÖNDORFF D & ROTHENPIELER UW (1999) Re-expression of the developmental gene Pax-2 during experimental acute tubular necrosis in mice 1. *Kidney Int.,* 56**,** 1423-31.

INNES BA & DORIN JR (2001) Submucosal gland distribution in the mouse has a genetic determination localized on chromosome 9. *Mamm Genome,* 12**,** 124-8.

IZBICKI G, SEGEL MJ, CHRISTENSEN TG, CONNER MW & BREUER R (2002) Time course of bleomycin-induced lung fibrosis. *Int J Exp Pathol. ,* 83**,** 111-119.

JEFFERY PK (2000) Comparison of the structural and inflammatory features of COPD and asthma. Giles F. Filley Lecture. *Chest,* 117**,** 251S-60S.

JEFFERY PK (2001) Remodeling in asthma and chronic obstructive lung disease. *Am J Respir Crit Care Med. ,* 164**,** S28-38.

JEFFERY PK (2004) Remodeling and inflammation of bronchi in asthma and chronic obstructive pulmonary disease. *Proc Am Thorac Soc,* 1**,** 176-183.

JEFFERY TK, UPTON PD, TREMBATH RC & MORRELL NW (2005) BMP4 inhibits proliferation and promotes myocyte differentiation of lung fibroblasts via Smad1 and JNK pathways. *Am J Physiol Lung Cell Mol Physiol,* 288**,** L370-8.

JEYASEELAN S, CHU HW, YOUNG SK & WORTHEN GS (2004) Transcriptional profiling of lipopolysaccharide-induced acute lung injury. *Infect Immun,* 72**,** 7247-7256.

KALLURI R & NEILSON EG (2003) Epithelial-mesenchymal transition and its implications for fibrosis. *J Clin Invest,* 112**,** 1776-84.

KEATING DT, MALIZIA AP, SADLIER D, HURSON C, WOOD AE, MCCARTHY J, NOLKE L, EGAN JJ & DORAN PP (2008) Lung tissue storage: optimizing conditions for future use in molecular research. *Exp Lung Res. ,* 34**,** 455-66.

KEENAN KP, COMBS JW & MCDOWELL EM (1982a) Regeneration of hamster tracheal epithelium after mechanical injury. I. Focal lesions: quantitative morphologic study of cell proliferation. *Virchows Arch B Cell Pathol Incl Mol Pathol,* 41**,** 193-214.

KEENAN KP, COMBS JW & MCDOWELL EM (1982b) Regeneration of hamster tracheal epithelium after mechanical injury. II. Multifocal lesions: stathmokinetic and autoradiographic studies of cell proliferation. *Virchows Arch B Cell Pathol Incl Mol Pathol.,* 41**,** 215-29.

KEENAN KP, COMBS JW & MCDOWELL EM (1982c) Regeneration of hamster tracheal epithelium after mechanical injury. III. Large and small lesions: comparative stathmokinetic and single pulse and continuous thymidine labeling autoradiographic studies. *Virchows Arch B Cell Pathol Incl Mol Pathol.,* 41**,** 231-52.

KEENAN KP, WILSON TS & MCDOWELL EM (1983) Regeneration of hamster tracheal epithelium after mechanical injury. IV. Histochemical, immunocytochemical and ultrastructural studies. *Virchows Arch B Cell Pathol Incl Mol Pathol,* 43**,** 213-240.

KEIR S & PAGE C (2008) The rabbit as a model to study asthma and other lung diseases. *Pulm Pharmacol Ther,* 21**,** 721-30.

KENNEDY AR, DESROSIERS A, TERZAGHI M & LITTLE JB (1978) Morphometric and histological analysis of the lungs of Syrian golden hamsters. *J Anat. ,* 125**,** 527-53.

KHATRI P & DRĂGHICI S (2005) Ontological analysis of gene expression data: current tools, limitations, and open problems. *Bioinformatics. ,* 21**,** 3587-3595.

KHO AT, BHATTACHARYA S, MECHAM BH, HONG J, KOHANE IS & MARIANI TJ (2009) Expression Profiles of the Mouse Lung Identify a Molecular Signature of Time-to-Birth *Am. J. Respir. Cell Mol. Bio.,* 40**,** 47-57.

KHO AT, BHATTACHARYA S, TANTISIRA KG, CAREY VJ, GAEDIGK R, LEEDER JS, KOHANE IS, WEISS ST & MARIANI TJ (2010) Transcriptomic analysis of human lung development. *Am J Respir Crit Care Med. ,* 181**,** 54-63.

KIM JS, MCKINNIS VS, ADAMS K & WHITE SR (1997a) Proliferation and repair of guinea pig tracheal epithelium after neuropeptide depletion and injury in vivo. *Am J Physiol,* 273**,** 1235-41.

KIM JS, MCKINNIS VS & WHITE SR (1997b) Migration of guinea pig airway epithelial cells in response to bombesin analogues. *Am J Respir Cell Mol Biol,* 16**,** 259-66.

KIM S & NADEL JA (2004) Role of neutrophils in mucus hypersecretion in COPD and implications for therapy. *Treat Respir Med,* 3**,** 147-59.

KIMURA J & DEUTSCH GH (2007) Key mechanisms of early lung development. *Pediatr Dev Pathol. ,* 10**,** 335-47.

KIRSCHVINK N & REINHOLD P (2008) Use of alternative animals as asthma models. *Curr Drug Targets,* 9**,** 470-484.

KOEMPEL JA, GIBSON SE, O'GRADY K & TORIUMI DM (1998) The effect of platelet-derived growth factor on tracheal wound healing. *Int J Pediatr Otorhinolaryngol,* 46**,** 1-8.

KOTTON DN, MA BY, CARDOSO WV, SANDERSON EA, SUMMER RS, WILLIAMS MC & FINE A (2001) Bone marrow-derived cells as progenitors of lung alveolar epithelium. *Development and disease,* 128 5181-5188.

KUMAR RK & FOSTER PS (2001) Murine model of chronic human asthma. *Immunol Cell Biol,* 79**,** 141-4.

LANE BP & GORDON R (1974) Regeneration of rat tracheal epithelium after mechanical injury. I. The relationship between mitotic activity and cellular differentiation. *Proc Soc Exp Biol Med* 145**,** 1139-1144.

LANE BP & GORDON RE (1979) Regeneration of vitamin A deficient rat tracheal epithelium after mild mechanical injury. Growth kinetics and cellular differentiation. *Differentiation,* 14**,** 87-93.

LAZARD DS, MOORE A, HUPERTAN V, MARTIN C, ESCABASSE V, DREYFUS P, BURGEL PR, AMSELEM S, ESCUDIER E & COSTE A (2009) Muco-ciliary differentiation of nasal epithelial cells is decreased after wound healing in vitro. *Allergy* 64**,** 1136-43.

LEE JM, DEDHAR S, KALLURI R & THOMPSON EW (2006) The epithelial-mesenchymal transition: new insights in signaling, development, and disease. *J Cell Biol,* 172**,** 973-981.

LESIMPLE P, VAN SEUNINGEN I, BUISINE MP, COPIN MC, HINZ M, HOFFMANN W, HAJJ R, BRODY SL, CORAUX C & PUCHELLE E (2007) Trefoil factor family 3 peptide promotes human airway epithelial ciliated cell differentiation. *Am J Respir Cell Mol Biol,* 36**,** 296-303.

LI C, XIAO J, HORMI K, BOROK Z & MINOO P (2002) Wnt5a participates in distal lung morphogenesis. *Dev Biol. ,* 248**,** 68-81.

LIU HW, CHENG B, LI JF, WU HJ, LI KY, SUN TZ & FU XB (2009) Characterization of angiotensin-converting enzyme expression during epidermis morphogenesis in humans: a potential marker for epidermal stem cells. *Br J Dermatol. ,* 160**,** 250-258.

LIU JY, NETTESHEIM P & RANDEL SH (1994) Growth and differentiation of tracheal epithelial progenitor cells. *Am. J. Physiol. Lung Cell Mol Physiol,* 266**,** 296-307.

LIU X, DRISKELL RR & ENGELHARDT JF (2006) Stem cells in the lung. *Methods Enzymol,* 419**,** 285-321.

LIU X & ENGELHARDT JF (2008) The glandular stem/progenitor cell niche in airway development and repair. *Proc Am Thorac Soc. ,* 5**,** 682-8.

MAEDA Y, DAVÉ V & WHITSETT JA (2007) Transcriptional control of lung morphogenesis. *Physiol Rev.,* 87**,** 219-44.

MARIANI TJ, REED JJ & SHAPIRO SD (2002) Expression profiling of the developing mouse lung: insights into the establishment of the extracellular matrix. *Am J Respir Cell Mol Biol. ,* 26**,** 541-8.

MARIASSY AT & PLOPPER CG (1983) Tracheobronchial epithelium of the sheep: I. Quantitative light-microscopic study of epithelial cell abundance, and distribution. *Anat Rec,* 205**,** 263-275.

MARTIN FJ & PRINCE AS (2008) TLR2 regulates gap junction intercellular communication in airway cells. *J Immunol. ,* 180**,** 4986-4993.

MCDOWELL EM, BECCI PJ, SCHÜRCH W & TRUMP BF (1979) The respiratory epithelium. VII. Epidermoid metaplasia of hamster tracheal epithelium during regeneration following mechanical injury. *J Natl Cancer Inst,* 62**,** 995-1008.

MCDOWELL EM, BEN T, NEWKIRK C, CHANG S & DE LUCA LM (1987) Differentiation of tracheal mucociliary epithelium in primary cell culture recapitulates normal fetal development and regeneration following injury in hamsters. *Am J Pathol,* 129**,** 511-522.

MCLACHLAN G, BAKER A, TENNANT P, GORDON C, VRETTOU C, RENWICK L, BLUNDELL R, CHENG SH, SCHEULE RK, DAVIES L, PAINTER H, COLES RL, LAWTON AE, MARRIOTT C, GILL DR, HYDE SC, GRIESENBACH U, ALTON EW, BOYD AC, PORTEOUS DJ &

COLLIE DD (2007) Optimizing aerosol gene delivery and expression in the ovine lung. *Mol Ther,* 15**,** 348-54.

MCLAUGHLIN RF, TYLER WS & CANADA RO (1994) A study of the subgross pulmonary anatomy in various mammals. *Am J Anat* 108 149 - 165.

MCNEIL JS, TORRINGTON KG, MUNDIE TG, BANKS RA, PHILLIPS YY & RIPPLE GR (1991) Prediction of carbon monoxide diffusing capacity of the lung in splenectomized sheep. *Lab Anim Sci,* 41**,** 63-5.

MENG QR, GIDEON KM, HARBO SJ, RENNE RA, LEE MK, BRYS AM & JONES R (2006) Gene expression profiling in lung tissues from mice exposed to cigarette smoke, lipopolysaccharide, or smoke plus lipopolysaccharide by inhalation. *Inhal Toxicol,* 18**,** 555-68.

MEYRICK B (1986) Pathology of the adult respiratory distress syndrome. *Crit Care Clin,* 2**,** 405-428.

MICKE P, OHSHIMA M, TAHMASEBPOOR S, REN ZP, OSTMAN A, PONTÉN F & BOTLING J (2006) Biobanking of fresh frozen tissue: RNA is stable in nonfixed surgical specimens. *Lab Invest. ,* 86**,** 202-11.

MILLER YE, BLATCHFORD P, HYUN DS, KEITH RL, KENNEDY TC, WOLF H, BYERS T, BUNN PA JR, LEWIS MT, FRANKLIN WA, HIRSCH FR & KITTELSON J (2007) Bronchial epithelial Ki-67 index is related to histology, smoking, and gender, but not lung cancer or chronic obstructive pulmonary disease. *Cancer Epidemiol Biomarkers Prev,* 16**,** 2425-31.

MONTESEIRÍN J (2009) Neutrophils and asthma. *J Investig Allergol Clin Immunol.,* 19**,** 340-354.

MUCENSKI ML, NATION JM, THITOFF AR, BESNARD V, XU Y, WERT SE, HARADA N, TAKETO MM, STAHLMAN MT & WHITSETT JA (2005) Beta-catenin regulates differentiation of respiratory epithelial cells in vivo. *Am J Physiol Lung Cell Mol Physiol,* 289**,** 971-979.

MUCENSKI ML, WERT SE, NATION JM, LOUDY DE, HUELSKEN J, BIRCHMEIER W, MORRISEY EE & WHITSETT JA (2003) Beta-catenin is required for specification of proximal/distal cell fate during lung morphogenesis. *J Biol Chem,* 278**,** 40231-8.

MUELLER O, LIGHTFOOT S & SCHROEDER A (2004) RNA integrity number (RIN)—standardization of RNA quality control. *Agilent Application Note,* Publication 5989-1165EN**,** 1–8.

MUNDIE TG, JANUSZKIEWICZ AJ, RAYBURN DB, MARTIN DG & RIPPLE GR (1991) Effects of conditioning and maximal incremental exercise on oxygen consumption in sheep. *Am J Vet Res,* 52**,** 1019-23.

MURPHY D (2002) Gene expression studies using microarrays: principles, problems, and prospects. *Adv Physiol Educ,* 26**,** 256-70.

NEURINGER IP & RANDELL SH (2004) Stem cells and repair of lung injuries. *Respir Res.,* 5**,** 6.

NEVEU WA, ALLARD JB, DIENZ O, WARGO MJ, CILIBERTO G, WHITTAKER LA & RINCON M (2009) IL-6 is required for airway mucus production induced by inhaled fungal allergens. *J Immunol. ,* 183**,** 1732-8.

NUSSE R, FUERER C, CHING W, HARNISH K, LOGAN C, ZENG A, TEN BERGE D & KALANI Y (2008) Wnt signaling and stem cell control. *Cold Spring Harb Symp Quant Biol.,* 73**,** 59-66.

OLVER RE & ROBINSON EJ (1986) Sodium and chloride transport by the tracheal epithelium of fetal, new-born and adult sheep. *J Physiol* 375**,** 377-390.

OTTO, W. R. (2002) Lung epithelial stem cells. *Journal of Pathology,* 197**,** 527–535.

OTTO WR (2002) Lung epithelial stem cells. *Journal of Pathology,* 197**,** 527–535.

PARK JW, RYTER SW & CHOI AM (2007) Functional significance of apoptosis in chronic obstructive pulmonary disease. *COPD,* 4**,** 347-53.

PARK KS, WELLS JM, ZORN AM, WERT SE, LAUBACH VE, FERNANDEZ LG & WHITSETT JA (2006) Transdifferentiation of ciliated cells during repair of the respiratory epithelium. *Am J Respir Cell Mol Biol.,* 34 151-157.

PATEL GK, WILSON CH, HARDING KG, FINLAY AY & BOWDEN PE (2006) Numerous keratinocyte subtypes involved in wound re-epithelialization. *J Invest Dermatol,* 126**,** 497-502.

PATERSON S & LELLO J (2003) Mixed models: getting the best use of parasitological data. *Trends Parasitol,* 19**,** 370-375.

PATTERSON, H. D. & THOMPSON, R. (1971) Recovery of inter-block information when block sizes are unequal. *Biometrika,* 58**,** 545-554.

PAYNE, R. W., HARDING, S. A., MURRAY, D. A., SOUTAR, D. M., BAIRD, D. B., GLASER, A. I., CHANNING, I. C., WELHAM, S. J., GILMOUR, A. R., THOMPSON, R. & WEBSTER, R. (2008) *GenStat Release 11 Reference Manual*, VSN International, Hemel Hempstead.

PILEWSKI JM, LATOCHE JD, ARCASOY SM & ALBELDA SM (1997) Expression of integrin cell adhesion receptors during human airway epithelial repair in vivo. *Am J Physiol. ,* 273**,** L256-63.

PINES J & HUNTER T (1992) Cyclins A and B1 in the human cell cycle. *Ciba Found Symp,* 170**,** 187-96.

PLOPPER CG & HYDE DM (2008) The non-human primate as a model for studying COPD and asthma. *Pulm Pharmacol Ther,* 21**,** 755-766.

PLOPPER CG, SUVERKROPP C, MORIN D, NISHIO S & BUCKPITT A (1992) Relationship of cytochrome P-450 activity to Clara cell cytotoxicity. I. Histopathologic comparison of the respiratory tract of mice, rats and hamsters after parenteral administration of naphthalene. *J Pharmacol Exp Ther.,* 261**,** 353-63.

PUCHELLE E, ZAHM JM, TOURNIER JM & CORAUX C (2006) Airway epithelial repair, regeneration, and remodeling after injury in chronic obstructive pulmonary disease. *Proc Am Thorac Soc,* 3**,** 726-733.

PUCHEU-HASTON CM, COPELAND LB, VALLANAT B, BOYKIN E & WARD MD (2010) Biomarkers of acute respiratory allergen exposure: Screening for sensitization potential. *Toxicol Appl Pharmacol.,* In-press.

QUARANTA V (1990) Epithelial integrins. *Cell Differ Dev. ,* 32**,** 361-5.

RAISER DM, ZACHAREK SJ, ROACH RR, CURTIS SJ, SINKEVICIUS KW, GLUDISH DW & KIM CF (2008) Stem cell biology in the lung and lung cancers: using pulmonary context and classic approaches. *Cold Spring Harb Symp Quant Biol,* 73**,** 479-90.

RAMAN T, O'CONNOR TP, HACKETT NR, WANG W, HARVEY BG, ATTIYEH MA, DANG DT, TEATER M & CRYSTAL RG (2009) Quality control in microarray assessment of gene expression in human airway epithelium. *BMC Genomics. ,* 10**,** 493.

RAPHAEL GD, JENEY EV, BARANIUK JN, KIM I, MEREDITH SD & KALINER MA (1989) Pathophysiology of rhinitis. Lactoferrin and lysozyme in nasal secretions. *J Clin Invest,* 84**,** 1528-35.

RAWLINS EL, OKUBO T, QUE J, XUE Y, CLARK C, LUO X & HOGAN BL (2008) Epithelial stem/progenitor cells in lung postnatal growth, maintenance, and repair. *Cold Spring Harb Symp Quant Biol. ,* 73.

RAWLINS EL, OSTROWSKI LE, RANDELL SH & HOGAN BLM (2007) Lung development and repair: Contribution of the ciliated lineage. *Proc Natl Acad Sci U S A. ,* 104 410–417.

RENNARD SI, BECKMAN J, DAUGHTON D, ERTL RF, KOYAMA S, RICKARD K, ROMBERGER D, RUBENSTEIN I, SHOJI S, SISSON J, SPURZEN J, THOMPSON AB, VON ESSEN S & ROBBINS RA (1991) *The role of airway epithelium in cellular migration,* New York, USA, Marcel Dekker, INC: Lung Biology in Health and Disease.

REYNOLDS SD, GIANGRECO A, POWER JH & STRIPP BR (2000) Neuroepithelial bodies of pulmonary airways serve as a reservoir of progenitor cells capable of epithelial regeneration. *Am J Pathol,* 156**,** 269-278.

REYNOLDS SD, ZEMKE AC, GIANGRECO A, BROCKWAYA BL, TEISANU RM, DRAKE JA, MARIANI T, DI PY, TAKETO MM & STRIPP BR (2008) Conditional stabilization of β-catenin expands the pool of lung stem cells. *Stem Cells* 26**,** 1337 -1346.

ROBERTS L, BOWERS J, SENSINGER K, LISOWSKI A, GETTS R & ANDERSON MG (2009) Identification of methods for use of formalin-fixed, paraffin-embedded tissue samples in RNA expression profiling. *Genomics. ,* 94**,** 341-348.

ROCK JR, ONAITIS MW, RAWLINS EL, LU Y, CLARK CP, XUE Y, RANDELL SH & HOGAN BL (2009) Basal cells as stem cells of the mouse trachea and human airway epithelium. *Proc Natl Acad Sci U S A. ,* 106**,** 12771-12775.

ROHRBECK A, NEUKIRCHEN J, ROSSKOPF M, PARDILLOS GG, GEDDERT H, SCHWALEN A, GABBERT HE, VON HAESELER A, PITSCHKE G,

SCHOTT M, KRONENWETT R, HAAS R & ROHR UP (2008) Gene expression profiling for molecular distinction and characterization of laser captured primary lung cancers. *J Transl Med* 6.

ROMAGNOLI M, VACHIER I, VIGNOLA AM, GODARD P, BOUSQUET J & CHANEZ P (1999) Safety and cellular assessment of bronchial brushing in airway diseases. *Respir Med. ,* 93**,** 461-6.

ROWE RK, BRODY SL & PEKOSZ A (2004) Differentiated cultures of primary hamster tracheal airway epithelial cells. *In Vitro Cell Dev Biol Anim. ,* 40**,** 303-11.

SAMBROOK J, FRITSCH EF & MANIATIS T (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, NY.

SCHEERLINCK JP, SNIBSON KJ, BOWLES VM & SUTTON P (2008) Biomedical applications of sheep models: from asthma to vaccines. *Trends Biotechnol. ,* 26**,** 259-66.

SCHITTNY, J. & YURCHENCO, P. (1989) Basement membranes: Molecular organisation and function in development and disease. *Curr Opin Cell Biol,* 1**,** 983-988.

SCHLOSSER RJ (2006) Surfactant and its role in chronic sinusitis. *Ann Otol Rhinol Laryngol Suppl,* 196**,** 40-4.

SCHROEDER A, MUELLER O, STOCKER S, SALOWSKY R, LEIBER M, GASSMANN M, LIGHTFOOT S, MENZEL W, GRANZOW M & RAGG T (2006) The RIN: an RNA integrity number for assigning integrity values to RNA measurements. *BMC Mol Biol. ,* 7**,** 3.

SHA Q, TRUONG-TRAN AQ, PLITT JR, BECK LA & SCHLEIMER RP (2004) Activation of airway epithelial cells by toll-like receptor agonists. *Am J Respir Cell Mol Biol. ,* 31**,** 358-364.

SHAFFER, J. (1995) Multiple hypothesis testing. *Ann Rev. Psychol.,* 46**,** 561-584.

SHANNON MJ & HYATT BA (2004) Epithelial mesenchymal interactions in the developing lung. *Annu. Rev. Physiol,* 66**,** 625-645.

SHI W, BELLUSCI S & WARBURTON D (2007) Lung development and adult lung diseases. *Chest,* 132**,** 651-6.

SHI W, CHEN F & CARDOSO WV (2009) Mechanisms of lung development: contribution to adult lung disease and relevance to chronic obstructive pulmonary disease. *Proc Am Thorac Soc. ,* 6**,** 558-63.

SHIM JJ, DABBAGH K, UEKI IF, DAO-PICK T, BURGEL PR, TAKEYAMA K, TAM DC & NADEL JA (2001) IL-13 induces mucin production by stimulating epidermal growth factor receptors and by activating neutrophils. *Am J Physiol Lung Cell Mol Physiol,* 280**,** 134-40.

SHIMIZU T, NISHIHARA M, KAWAGUCHI S & SAKAKURA Y (1994) Expression of phenotypic markers during regeneration of rat tracheal epithelium following mechanical injury. *Am J Respir Cell Mol Biol,* 11**,** 85-94.

SHULTZ MA, ZHANG L, GU YZ, BAKER GL, FANNUCHI MV, PADUA AM, GURSKE WA, MORIN D, PENN SG, JOVANOVICH SB, PLOPPER CG & AR, B. (2004) Gene expression analysis in response to lung toxicants: I. Sequencing and microarray development. *Am J Respir Cell Mol Biol,* 30**,** 296-310.

SKAWRAN B, DIERICH M, STEINEMANN D, HOHLFELD J, HAVERICH A, SCHLEGELBERGER B, WELTE T & VON NEUHOFF N (2009) Bronchial epithelial cells as a new source for differential transcriptome analysis after lung transplantation. *Eur J Cardiothorac Surg. ,* 36**,** 715-21.

SNYDER JC, ZEMKE AC & STRIPP BR (2009) Reparative capacity of airway epithelium impacts deposition and remodeling of extracellular matrix. *Am J Respir Cell Mol Biol. ,* 40**,** 633-42.

SRINATH REDDY K, SHAH B, VARGHESE C & RAMADOSS A (2005) Responding to the threat of chronic diseases in India. *Lancet,* 366**,** 1744-1749.

STEILING H & WERNER S (2003) Fibroblast growth factors: key players in epithelial morphogenesis, repair and cytoprotection. *Curr Opin Biotechnol. ,* 14**,** 533-537.

STEWART, C. (2006) *Stem cells and stem cell therapies: The future for regenerative medicine?,* New York, Nova Science Publishers, Inc. pp 119-142.

STICK S (2000) Pediatric origins of adult lung disease. 1. The contribution of airway development to paediatric and adult lung disease. *Thorax,* 55**,** 587-94.

STRIPP BR & REYNOLDS SD (2008) Maintenance and repair of the bronchiolar epithelium. *Proc Am Thorac Soc. ,* 5**,** 328-33.

STRIPP, B. R. & SHAPIRO, S. D. (2006) Stem Cells in Lung Disease, Repair, and the Potential for Therapeutic Interventions: State-of-the-Art and Future Challenges. *Am J Respir Cell Mol Biol,* 34**,** 517–522.

SUE SR, CHARI RS, KONG FM, MILLS JJ, FINE RL, JIRTLE RL & MEYERS WC (1995) Transforming growth factor-beta receptors and mannose 6-phosphate/insulin-like growth factor-II receptor expression in human hepatocellular carcinoma. *Ann Surg. ,* 222**,** 171-178.

SZILASI M, DOLINAY T, NEMES Z & STRAUSZ J (2006) Pathology of chronic obstructive pulmonary disease. *Pathol Oncol Res,* 12**,** 52-60.

TAKAHASHI K & COULOMBE PA (1997) Defining a region of the human keratin 6a gene that confers inducible expression in stratified epithelia of transgenic mice. *J Biol Chem. ,* 272**,** 11979-85.

TEBBUTT SJ, HARRIS A & HILL DF (1996) An ovine CFTR variant as a putative cystic fibrosis causing mutation. *J Med Genet,* 33**,** 623-4.

TEBBUTT SJ, WARDLE CJ, HILL DF & HARRIS, A. (1995) Molecular analysis of the ovine cystic fibrosis transmembrane conductance regulator gene. *Proc Natl Acad Sci U S A,* 92**,** 2293-2297.

TESFAIGZI Y (2003) Processes involved in the repair of injured airway epithelia. *Arch Immunol Ther Exp (Warsz). ,* 51**,** 283-8.

THEISE ND, HENEGARIU O, GROVE J, JAGIRDAR J, KAO PN, CRAWFORD JM, BADVE S, SAXENA R & DS, K. (2002) Radiation pneumonitis in mice: a severe injury model for pneumocyte engraftment from bone marrow. *Exp Hematol. ,* 30**,** 1333-1338.

THEOCHARIDIS A, VAN DONGEN S, ENRIGHT AJ & FREEMAN TC (2009) Network visualization and analysis of gene expression data using BioLayout Express(3D). *Nat Protoc.,* 4**,** 1535-1550.

TÖLGYESI G, MOLNÁR V, SEMSEI AF, KISZEL P, UNGVÁRI I, PÓCZA P, WIENER Z, KOMLÓSI ZI, KUNOS L, GÁLFFY G, LOSONCZY G, SERES I, FALUS A & SZALAI C (2009) Gene expression profiling of experimental asthma reveals a possible role of paraoxonase-1 in the disease. *Int Immunol. ,* 21**,** 967-975.

TUCK SA, RAMOS-BARBÓN D, CAMPBELL H, MCGOVERN T, KARMOUTY-QUINTANA H & MARTIN JG (2008) Time course of airway remodelling after an acute chlorine gas exposure in mice. *Respir Res. ,* 9**,** 61.

VAN GELDER RN, VON ZASTROW ME, YOOL A, DEMENT WC, BARCHAS JD & EBERWINE JH (1990) Amplified RNA synthesized from limited quantities of heterogeneous cDNA. *Proc Natl Acad Sci U S A. ,* 87**,** 1663-7.

VORTKAMP A, PATHI S, PERETTI GM, CARUSO EM, ZALESKE DJ & TABIN CJ (1998) Recapitulation of signals regulating embryonic bone formation during postnatal growth and in fracture repair. *Mech Dev. ,* 71**,** 65-76.

VOYNOW JA, FISCHER BM, ROBERTS BC & PROIA AD (2005) Basal-like cells constitute the proliferating cell population in cystic fibrosis airways *Am J Respir Crit Care Med,* 172**,** 1013-1018.

WAGNER EM & MITZNER WA (1990) Bronchial circulatory reversal of methacholine-induced airway constriction. *J Appl Physiol,* 69**,** 1220-4.

WALKER HL, MCLEOD CG JR & MCMANUS WF (1981) Experimental inhalation injury in the goat. *J Trauma. ,* 21**,** 962-964.

WALTHER FJ, WARING AJ, SHERMAN MA, ZASADZINSKI JA & GORDON LM (2007) Hydrophobic surfactant proteins and their analogues. *Neonatology,* 91**,** 303-10.

WANG CZ, LI A & YANG ZC (1990) The pathophysiology of carbon monoxide poisoning and acute respiratory failure in a sheep model with smoke inhalation injury. *Chest,* 97**,** 736-42.

WANG IM, STEPANIANTS S, BOIE Y, MORTIMER JR, KENNEDY B, ELLIOTT M, HAYASHI S, LOY L, COULTER S, CERVINO S, HARRIS J, THORNTON M, RAUBERTAS R, ROBERTS C, HOGG JC, CRACKOWER M, O'NEILL G & PARÉ PD (2008) Gene expression profiling in patients with chronic obstructive pulmonary disease and lung cancer. *Am J Respir Crit Care Med. ,* 177**,** 402-11.

WANG K, WEN FQ, FENG YL, OU XM, XU D, YANG J & DENG ZP (2007) Increased expression of human calcium-activated chloride channel 1 gene is correlated with mucus overproduction in Chinese asthmatic airway. *Cell Biol Int,* 31**,** 1388-1395.

WANG L, KONG L, WU F, BAI Y & BURTON R (2005) Preventing chronic diseases in China. *Lancet. ,* 366**,** 1821-1824.

WANG S, CHAN A, ILSLEY D & COATES S (2003) Quantitative analysis of sample RNA inputs on microarray performance using Agilent Fluorescent Direct Labeling kit. *Application note,* Agilent Technologies, Inc.

WANNER A, MEZEY RJ, REINHART ME & EYRE P (1979) Antigen-induced bronchospasm in conscious sheep. *J Appl Physiol,* 47**,** 917-22.

WANNER A & REINHART ME (1978) Respiratory mechanics in conscious sheep: response to methacholine. *J Appl Physiol,* 44**,** 479-82.

WARBURTON D, BELLUSCI S, DE LANGHE S, DEL MORAL PM, FLEURY V, MAILLEUX A, TEFFT D, UNBEKANDT M, WANG K & SHI W (2005) Molecular mechanisms of early lung specification and branching morphogenesis. *Pediatr Res,* 57**,** 26-37.

WARBURTON D, SCHWARZ M, TEFFT D, FLORES-DELGADO G, ANDERSON KD & CARDOSO WV (2000) The molecular basis of lung morphogenesis. *Mechanisms of Development,* 92**,** 55-81.

WARBURTON D, TEFFT D, MAILLEUX A, BELLUSCI S, THIERY JP, ZHAO J, BUCKLEY S, SHI W & DRISCOLL B (2001) Do lung remodeling, repair, and regeneration recapitulate respiratory ontogeny? *Am J Respir Crit Care Med. ,* 164**,** S59-62.

WARBURTON D, WUENSCHELL C, FLORES-DELGADO G & ANDERSON K (1998) Commitment and differentiation of lung cell lineages. *Biochem Cell Biol. ,* 76**,** 971-95.

WATT FM & HOGAN BL (2000) Out of Eden: stem cells and their niches. *Science,* 287**,** 1427-1430.

WERLEN C, PY P & HAAB P (1984) Alveolar-arterial equilibration in the lung of sheep. *Respir Physiol,* 55**,** 205-21.

WESLEY UV, BOVE PF, HRISTOVA M, MCCARTHY S & VAN DER VLIET A (2007) Airway epithelial cell migration and wound repair by ATP-mediated activation of dual oxidase 1. *J Biol Chem. ,* 282**,** 3213-3220.

WESTPHAL M, COX RA, TRABER LD, MORITA N, ENKHBAATAR P, SCHMALSTIEG FC, HAWKINS HK, MAYBAUER DM, MAYBAUER MO, MURAKAMI K, BURKE AS, WESTPHAL-VARGHESE BB, RUDLOFF HE, SALSBURY JR, JODOIN JM, LEE S & TRABER DL (2006) Combined burn and smoke inhalation injury impairs ovine hypoxic pulmonary vasoconstriction. *Crit Care Med,* 34**,** 1428-36.

WHEELER AP, HARDIE WD & BERNARD G (1990a) Studies of an antiendotoxin antibody in preventing the physiologic changes of endotoxemia in awake sheep. *Am Rev Respir Dis,* 142**,** 775-81.

WHEELER AP, HARDIE WD & BERNARD GR (1992) The role of cyclooxygenase products in lung injury induced by tumor necrosis factor in sheep. *Am Rev Respir Dis,* 145**,** 632-9.

WHEELER AP, JESMOK G & BRIGHAM KL (1990b) Tumor necrosis factor's effects on lung mechanics, gas exchange, and airway reactivity in sheep. *J Appl Physiol,* 68**,** 2542-9.

WHITE SR, DORSCHEID DR, RABE KF, WOJCIK KR & HAMANN KJ (1999) Role of very late adhesion integrins in mediating repair of human airway epithelial cell monolayers after mechanical injury. *Am J Respir Cell Mol Biol. ,* 20**,** 787-96.

WHO (2001) WHO consultation on the development of a comprehensive approach for the prevention and control of chronic respiratory diseases. *World Health Organization*.

WILHELM DL (1953) Regeneration of tracheal epithelium. *J Pathol Bacteriol,* 65**,** 543-50.

WOJCIK SM, BUNDMAN DS & ROOP DR (2000) Delayed wound healing in keratin 6a knockout mice. *Mol Cell Biol. ,* 20**,** 5248-55.

XIAO J, LUCAS A, D'ANDRADE P, VISITACION M, TANGVORANUNTAKUL P & FULMER-SMENTEK S (2006) Performance of the Agilent microarray platform for one-colour analysis of gene expression: Application note. *Agilent Technologies, Inc.*

YANAI T, MATSUMOTO C, TAKASHIMA H, YOSHIDA K, SAKAI H, ISOWA K, IWASAKI T, SATO Y & MASEGI T (1996) Immunohistochemical demonstration of S-phase cells by anti-bromodeoxyuridine monoclonal antibody in cattle tissues. *J Comp Pathol.,* 114**,** 265-272.

YINGJIAN YOU, TAO HUANG, EDWARD J. RICHER, JENS-ERIK HARBOE SCHMIDT, JOSEPH ZABNER, ZEA BOROK & BRODY, S. L. (2004) Role of f-box factor foxj1 in differentiation of ciliated airway epithelial cells. *Am J Physiol Lung Cell Mol Physiol,* 286**,** 650–657.

YONEDA K, CHANG MM, CHMIEL K, CHEN Y & WU R (2003) Application of high-density DNA microarray to study smoke- and hydrogen peroxide-induced injury and repair in human bronchial epithelial cells. *J Am Soc Nephrol,* 14**,** 284-289.

YOO J, GHIASSI M, JIRMANOVA L, BALLIET AG, HOFFMAN B, FORNACE AJ JR, LIEBERMANN DA, BOTTINGER EP & ROBERTS AB (2003) Transforming growth factor-beta-induced apoptosis is mediated by Smad-dependent expression of GADD45b through p38 activation. *J Biol Chem. ,* 278**,** 43001-43007.

YOU Y, HUANG T, RICHER EJ, SCHMIDT JE, ZABNER J, BOROK Z & BRODY SL (2004) Role of f-box factor foxj1 in differentiation of ciliated airway epithelial cells. *Am J Physiol Lung Cell Mol Physiol,* 286**,** 650-657.

ZAHM JM, CHEVILLARD M & PUCHELLE E (1991) Wound repair of human surface respiratory epithelium. *Am J Respir Cell Mol Biol,* 5**,** 242-8.

ZAVADIL J & BÖTTINGER EP (2005) TGF-beta and epithelial-to-mesenchymal transitions. *Oncogene,* 24**,** 5764-5774.

ZEMKE AC, TEISANU RM, GIANGRECO A, DRAKE JA, BROCKWAY BL, REYNOLDS SD & STRIPP BR (2009) Beta-catenin is not necessary for maintenance or repair of the bronchiolar epithelium. *Am J Respir Cell Mol Biol.,* 41**,** 535-543.

ZHANG Y, GOSS AM, COHEN ED, KADZIK R, LEPORE JJ, MUTHUKUMARASWAMY K, YANG J, DEMAYO FJ, WHITSETT JA, PARMACEK MS & MORRISEY EE (2008) A Gata6-Wnt pathway required for epithelial stem cell development and airway regeneration. *Nat Genet. ,* 40**,** 862-70.

ZHAO P, H. E. (2004) Embryonic myogenesis pathways in muscle regeneration. *Dev Dyn. ,* 229**,** 380-392.

ZHUANG S, YAN Y, HAN J & SCHNELLMANN RG (2005) p38 kinase-mediated transactivation of the epidermal growth factor receptor is required for dedifferentiation of renal epithelial cells after oxidant injury. *J Biol Chem,* 280**,** 21036-21042.

APPENDIX

Sheep ID	Age (d)	Sheep Breed	Gender	Bodyweight (kg)
11934	435	Suff x	MN	49
11784	443	Suff x	MN	57
12223	456	Suff x	MN	56
7458	457	Suff x	F	50
11907	512	Suff x	MN	65
12171	517	Suff x	MN	52
10908	520	Suff x	MN	66
10639	526	Suff x	MN	75
77	530	Suff x	MN	65

Appendix 2.7.1: Details of the animals used in this study
Suff x = Suffolk crossbred, MN = Male neutered, F = female.

A)

SHEEP	TIMEPOINTS								
	-35 days	-28 days	-21 days	-14 days	-7 days	-3 days	-1 day	-6 hours	0 hour
11934	√	√	√	√	√	√	√	-	Necropsy
11784	√	√	√	√	√	√	√	-	
12223	-	√	√	√	√	√	√	√	
7458	√	√	√	√	√	√	√	-	

B)

SHEEP	TIMEPOINTS					
	-14 days	-7 Days	-3 days	-1 day	-6 hours	0 hour
217	-	-	-	-	√	Necropsy
10908	-	-	-	√	-	
77	-	-	√	-	-	
12171	-	√	-	-	-	
10639	√	-	-	-	-	

Appendix 2.7.2: Experimental protocols employed in this study
Appendix 2.7.2A depicts the sampling schedule employed in Protocol A and Appendix 2.7.2B, the sampling schedule employed in protocol B. In Protocol A, only one sheep (one segment) was evaluated 6 hours after injury (Sheep 12223)

SeqA	Name	Len(nt)	SeqB	Name	Len(nt)	Score
1	**CloneJ**	**403**	**2**	**TC3860**	**1654**	**94**
1	**CloneJ**	**403**	**3**	**CloneQ**	**403**	**79**
1	**CloneJ**	**403**	**4**	**42**	**1388**	**90**
1	**CloneJ**	**403**	**5**	**17**	**1568**	**96**
1	**CloneJ**	**403**	**6**	**13**	**1741**	**79**
1	**CloneJ**	**403**	**7**	**14**	**1771**	**92**
1	**CloneJ**	**403**	**8**	**15**	**1841**	**79**
2	TC3860	1654	3	CloneQ	403	79
2	TC3860	1654	4	42	1388	69
2	TC3860	1654	5	17	1568	69
2	TC3860	1654	6	13	1741	50
2	TC3860	1654	7	14	1771	97
2	TC3860	1654	8	15	1841	50
3	**CloneQ**	**403**	**4**	**42**	**1388**	**80**
3	**CloneQ**	**403**	**5**	**17**	**1568**	**78**
3	**CloneQ**	**403**	**6**	**13**	**1741**	**91**
3	**CloneQ**	**403**	**7**	**14**	**1771**	**78**
3	**CloneQ**	**403**	**8**	**15**	**1841**	**91**
4	42	1388	5	17	1568	77
4	42	1388	6	13	1741	62
4	42	1388	7	14	1771	69
4	42	1388	8	15	1841	62
5	17	1568	6	13	1741	59

Appendix 3.1: A summary of the alignment analysis result between clones and other types of bovine keratin genes. Highlighted (bolded) are the clones J and Q. The data shows that the sequence for Clone J is highly likely similar to bovine keratin 17, keratin 14 and EST sequence (TC3860) which is also has 97% similar to bovine keratin 14 (not shown). The sequence for Clone Q is likely to suggest that this sequence either belongs to bovine keratin 15 or 13.

SeqA : Numbering system assigned by the programme for the studied sequences
Name : The names given to the sequences
Len(nt) : The length of the analysed sequences
SeqB : Numbering system taken from SeqA
Name : The names given to the sequences
Len(nt) : The length of the analysed sequences
Score : The percentage of the sequence similarity as compared between sequences

```
TC3860          GGCCGACATCAACGGCCTGCGCAGGGTGCTGGACGAGCTGACCCTGGCCAGAGCCGACCT 771
14              GGCCGACATCAACGGCCTGCGTAGGGTGCTGGACGAGCTGACCCTGGCCAGAGCCGACCT 891
CloneJ          ----GACATCAACGGCCTGCGCAGGGTGCTGGACGAGCTGACCCTGGCCAGAGCTGACCT 56
17              GGCCGACATCAACGGCCTGCGCAGGGTGCTGGACGAGCTGACCCTGGCCAGAGCCGACCT 699
42              GGCCGACATCAACGGCCTGCGCAGGGTGCTGGACGAGCTGACCCTGGCCAGAGCCGACCT 670
13              GGCCGACATCAACGGCCTGCGCAGGGTGCTGGATGAGCTGACCCTGACCAAGACTGACCT 735
15              GGCCGACATCAACGGCCTGCGCAGGGTGCTGGATGAGCTGACCCTGACCAAGACTGACCT 865
CloneQ          ----GACATCAACGGCCTGCGCAGGGTGCTGGACGAGCTGACCCTGACCAAGACTGATCT 56
                    ***************** *********** ************ ***   * ** **

TC3860          GGAGATGCAGATCGAGAGCCTCAAGGAGGAGCTGGCCTACCTCCGGAAGAACCACGAGGA 831
14              GGAGATGCAGATCGAGAGCCTCAAGGAGGAGCTGGCCTACCTCCGGAAGAACCACGAGGA 951
CloneJ          GGAGATGCAGATCGAGAACCACTTGGAGGAGCTGGCCTACCTCCGCAAGAACCACGAGGA 116
17              GGAGATGCAGATCGAGAACCTCAAGGAGGAGCTGGCCTACCTCCGCAAGAATCACGAGGA 759
42              GGAGATGCAGATCGAGAACCTCAAGGAGGAGCTGGCCTACCTCCGGAAGAACCACGAGGA 730
13              GGAGATGCAGATCGAGAGCCTGAATGAGGAGCTGGCCTACCTGAAGAAGAACCACGAGGA 795
15              GGAGATGCAGATCGAGAGCCTGAATGAGGAGCTGGCCTACCTGAAGAAGAACCACGAGGA 925
CloneQ          GGAGATGCAGATTGAGAGCCTGAACGAGGAGCTGGCCTACCTGAAGAAGAACCACGAGGA 116
                ************ **** **     *****************    ***** ********

TC3860          GGAAATGAACTCCCTGAGAGGCCAGGTGGGCGGAGACGTCAACGTGGAGATGGACGCCGC 891
14              GGAAATGAACTCCCTGAGAGGCCAGGTGGGCGGAGACGTCAACGTGGAGATGGACGCCGC 1011
CloneJ          GGAGATGAAAGCCCTTCGAGGCCAGGTGGGTGGCGAGATCAACGTGGAGATGGACGCCGC 176
17              GGAGATGAAAGCCCTGCGAGGCCAGGTGGGCGGCGAGATCAACGTGGAGATGGACGCTGC 819
42              GGAGATGAATGCTCTGCGAGGCCAGGTGGGCGGGGACGTCAACGTGGAGATGGACGCCGC 790
13              GGAGATGAAGGAGTTCAGCAACCAGCTGGCTGGCCAGGTCAATGTGGAGATGGATGCAGC 855
15              GGAGATGAAGGAGTTCAGCAACCAGCTGGCTGGCCAGGTCAATGTGGAGATGGATGCAGC 985
CloneQ          GGAGATGAAGGAGTTCAGCAACCAGCTGGCTGGCCAGGTCAACGTGGAGATGGACGCAGC 176
                *** *****     *  *   **** ***  **  *  **** *********** ** **
```

```
TC3860          CCCCGGCGTGGACCTGAGCCGCATCCTGAACGAGATGCGCGACCAGTACGAGAAGATGGC 951
14              CCCCGGCGTGGACCTGAGCAGAATCCTGAACGAGATGCGCGACCAGTACGAGAAGATGGC 1071
CloneJ          CCCCGGCGTGGACCTGAGCCGCATCCTGAACGAGATGCGCGACCAGTATGAGAAGATGGC 236
17              CCCCGGCGTGGACCTGAGCCGCATCCTGAACGAGATGCGCGACCAGTATGAGAAGATGGC 879
42              CCCTGGCGTGGACCTGAGCCGCATCCTGAACGAGATGCGCGACCAGTATGAGAAGATGGC 850
13              ACCAGCCGTGGACCTGACCCGCGTGCTGTCAGAAATGAGGGAGCAGTACGAGGCCATGGC 915
15              ACCAGCCGTGGACCTGACCCGCGTGCTGTCAGAAATGAGGGAGCAGTACGAGGCCATGGC 1045
CloneQ          ACCAGGCATTGACCTGACCCGCGTGCTGGCAGAGATGAGGGAGCAGTACGAGGCCATGGC 236
                 ** * * * ******* * *  * ***   ** *** * ** ***** ***   *****

TC3860          AGAGAAGAACCGCAAGGATGCCGAGGACTGGTTCTTCAGCAAGACAGAGGAACTGAACCG 1011
14              GGAGAAGAACCGCAAGGATGCCGAAGACTGGTTCTTTAGCAAGACAGAGGAACTGAACCG 1131
CloneJ          AGAGAAGAACCGCAAGGATGCCGAGGACTGGTTCTTCAGCAAGACGGAGGAACTGAACCG 296
17              GGAGAAGAACCGCAAGGATGCCGAGGACTGGTTCTTTAGCAAGACGGAGGAACTGAACCG 939
42              GGAGAAGAACCGCAAGGATGCCGAGGACTGGTTCTTCAGCAAGACGGAGGAGCTGAACCG 910
13              GGAGAAGAACCGCCGGGATGCCGAGGCCTGGTTCTTTAGCAAGACAGAGGAGCTGAATAA 975
15              GGAGAAGAACCGCCGGGATGCCGAGGCCTGGTTCTTTAGCAAGACAGAGGAGCTGAATAA 1105
CloneQ          AGAGAAGAACCGCCGGGATGCCGAGGAATGGTTCCACAGCAAGAGCGCCGAGCTGAACAA 296
                 ************  ********* *  ******   *******  *  ** *****

TC3860          CGAGGTGGCTACCAACAGCGAGCTGGTGCAGAGCGGCAAGAGCGAGATCTCGGAGCTCCG 1071
14              CGAGGTGGCTACCAACAGCGAGCTGGTGCAGAGCGGCAAGAGCGAAATCTCGGAGCTCCG 1191
CloneJ          CGAGGTGGCCACCAACAGCGAGCTGGAGCAGAGCGGCAAGAGCGAGATCTCGGAGCTCCG 356
17              CGAGGTGGCCACCAACAGCGAGCTGGTGCAGAGCGGCAAGAGCGAGATCTCGGAGCTCCG 999
42              CGAGGTGGCCACCAACACCGAAGCCCTGCAGAGCAGCCGGACGGAGATCACGGAGCTGCG 970
13              GGAGGTGGCCTCCAACACAGAGATGATCCAGACCAGCAAGACGGAGATCACAGACCTGAG 1035
15              GGAGGTGGCCTCCAACACAGAGATGATCCAGACCAGCAAGACGGAGATCACAGACCTGAG 1165
CloneQ          GGAGGTGTCTTCCAGCACAGCCATGATCCAGACCAGCAAGACAGAGATCACCGAGCTCAG 356
                 ****** *  *** **  *        **** * **  **  ** *** * ** **  *
```

```
TC3860          GCGCACCATGCAAAACCTGGAGATCGAGCTGCAGTCCCAGCTCAGCATGAAAGCATCTCT 1131
14              GCGCACCTTGCAAAACCTGGAGATCGAGCTGCAGTCCCAGCTCAGCATGAAAGCATCTCT 1251
CloneJ          GCGCACCTTGCAGGCCCTGGAGATCGTGCTGCAGTCCCAGCTCAGCA-------------- 403
17              GCGCACCTTGCAGGCCCTGGAGATCGAACTGCAGTCCCAGCTCAGCATGAAAGCATCCCT 1059
42              CCGCACGGTCCAGAACCTGGAGATTGAGCTGCAGTCCCAGCTCAGCATGAAAGCGTCGCT 1030
13              ACGCACGATACAGGGGCTGGAGATTGAGCTGCAGTCCCAGCTCAGCATGAAAGCCGGGCT 1095
15              ACGCACGATACAGGGGCTGGAGATTGAGCTGCAGTCCCAGCTCAGCATGAAAGCCGGGCT 1225
CloneQ          GCGCACGCTCCAGGGCCTGGAGATCGAGCTGCAGTCCCAGCTCAGCA-------------- 403
                 *****  * **    ******** *  ******************
```

Appendix 3.2: Sequences alignment between the sequences generated from the Clones J and Q and other types of keratin genes. The analysis was performed in order to find a similarity of the sequences between clones and other types of keratin genes. Highlighted (bolded) sequences are the forward and reverse primers. The sequences from both forward and reverse primers are conserved between majorities of keratin genes. (*) shows the areas of conserved between studied sequences.